SEARS FOUNDATION FOR MARINE RESEARCH
BINGHAM OCEANOGRAPHIC LABORATORY, YALE UNIVERSITY

Journal of Marine Research

SVERDRUP
SIXTIETH ANNIVERSARY

VOLUME VII, 1948
NUMBER 3

Issued November 15, 1948, New Haven, Conn., U.S.A.

JOURNAL OF MARINE RESEARCH

PUBLISHED BY

SEARS FOUNDATION FOR MARINE RESEARCH
BINGHAM OCEANOGRAPHIC LABORATORY, YALE UNIVERSITY

BOARD OF EDITORS

Chairman: A. E. PARR
American Museum
of Natural History
New York

C. O'D. ISELIN
Woods Hole Oceanographic
Institution

H. U. SVERDRUP
Scripps Institution
of Oceanography

DANIEL MERRIMAN
Bingham Oceanographic
Laboratory

C.-G. ROSSBY
University of
Chicago

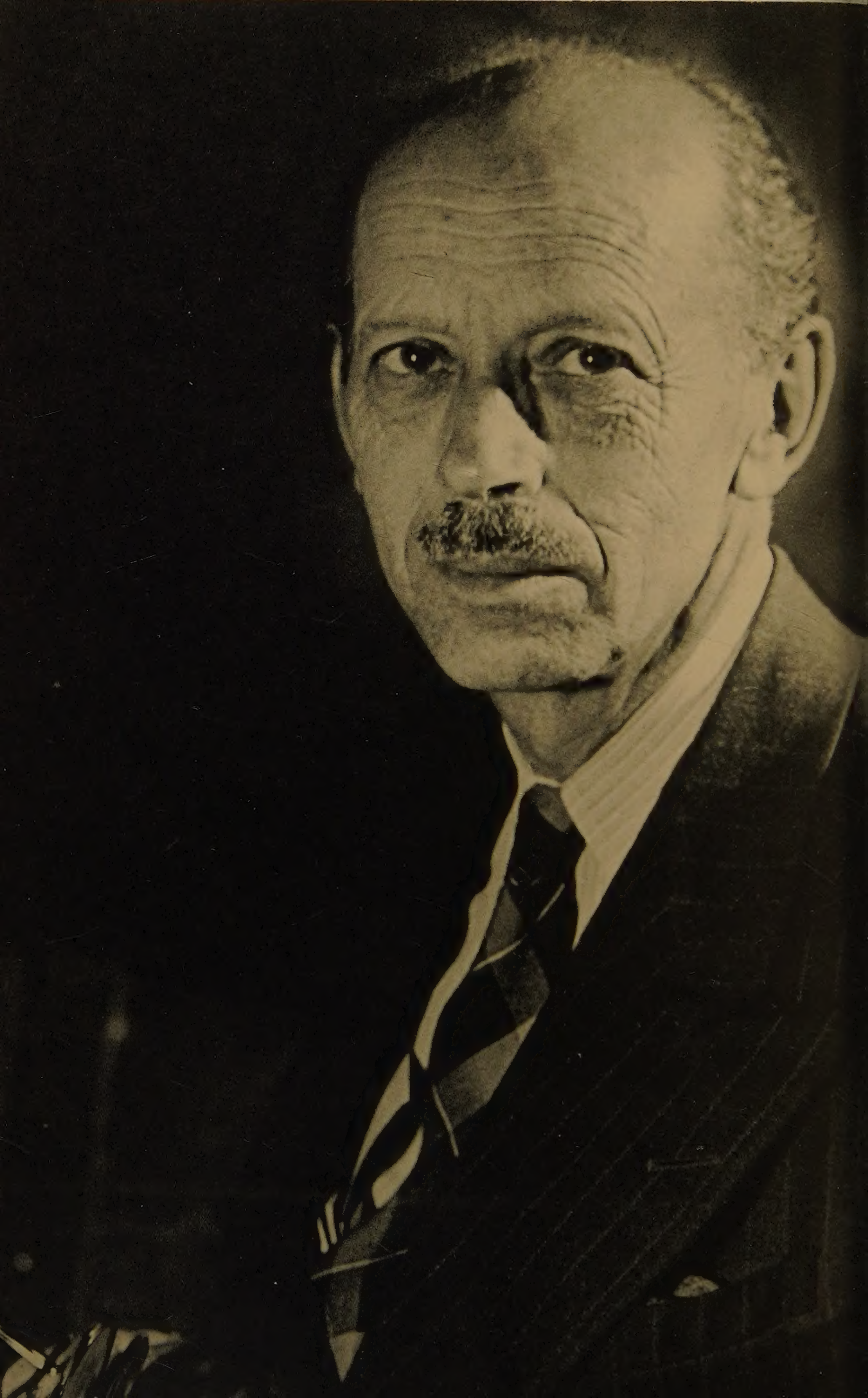
THOMAS G. THOMPSON
Oceanographic Laboratories
University of Washington

MANAGING EDITOR

YNGVE H. OLSEN
Bingham Oceanographic Laboratory
Yale University
New Haven, Conn.

The Journal is normally issued in three numbers per year constituting one volume, except for volume V, published in 1942–1944, and volume VI, published in 1946 and 1947. All the income derived from subscriptions is used solely to increase the volume and improve the quality of the Journal for the benefit of its subscribers and contributors.

Subscription price three dollars per volume, or one dollar per number. Address all communications to the Managing Editor. See inside rear cover for Note to Authors.



THIS NUMBER IS DEDICATED TO

HARALD ULRIK SVERDRUP

ON THE OCCASION OF

HIS SIXTIETH BIRTHDAY

BY SOME OF HIS FRIENDS AND ASSOCIATES IN AMERICA, IN APPRECIATION OF HIS PROFOUND INFLUENCE ON THE DEVELOPMENT OF OCEANOGRAPHY AND METEOROLOGY

The extra cost of this number of the JOURNAL OF MARINE RESEARCH has been made possible by a generous grant from the Ellen Browning Scripps Foundation.

←*Photograph by Paul Williams*

TABLE OF CONTENTS

PAGE

Harald Ulrik Sverdrup—An Appreciation 127

Bibliography of Harald Ulrik Sverdrup 132

1. **Daniel Merriman**: A posse ad esse 139
2. **J. A. Fleming**: Oceanography and geomagnetism 147
3. **Jacob Bjerknes**: Atmospheric tides 154
4. **Jörgen Holmboe**: On dynamic stability of zonal currents 163
5. **C.-G. Rossby**: On displacements and intensity changes of atmospheric vortices 175
6. **George F. McEwen**: The dynamics of large horizontal eddies (axes vertical) in the ocean off southern California 188
7. **Bernhard Haurwitz**: The effect of ocean currents on internal waves .. 217
8. **Victor P. Starr and George W. Platzman**: The transmission of energy by gravity waves of finite height 229
9. **J. D. Cochrane and R. S. Arthur**: Reflection of Tsunamis 239
10. **Theodore von Kármán**: Progress in the statistical theory of turbulence 252
11. **Carl Eckart**: An analysis of the stirring and mixing processes in incompressible fluids 265
12. **Walter H. Munk and Ernest R. Anderson**: Notes on a theory of the thermocline 276
13. **Donald W. Pritchard**: Streamlines from a discrete vector field: With application to ocean currents 296
14. **Robert O. Reid**: A model of the vertical structure of mass in equatorial wind-driven currents of a baroclinic ocean 304
15. **Pier Groen**: Methods for estimating dynamic slopes and currents in shallow water 313
16. **C. O'D. Iselin and Frederick C. Fuglister**: Some recent developments in the study of the Gulf Stream 317
17. **Woodrow C. Jacobs**: Some empirical relations between seasonal values of $(E-P)$ and surface salinities over the North Atlantic 330
18. **Dale F. Leipper**: Fog development at San Diego, California 337
19. **Alfred C. Redfield**: The exchange of oxygen across the sea surface .. 347
20. **Norris W. Rakestraw and Dayton E. Carritt**: Some seasonal chemical changes in the open ocean 362
21. **Athelstan F. Spilhaus and Arthur R. Miller**: The sea sampler. .. 370

PAGE

22. Henry Stommel: The theory of the electric field induced in deep ocean currents.......... 386
23. Russell W. Raitt: Sound scatterers in the sea.......... 393
24. V. O. Knudsen, R. S. Alford and J. W. Emling: Underwater ambient noise.......... 410
25. Robert S. Dietz: Deep scattering layer in the Pacific and Antarctic oceans.......... 430
26. Martin W. Johnson: Sound as a tool in marine ecology, from data on biological noises and the deep scattering layer.......... 443
27. Carl L. Hubbs: Changes in the fish fauna of western North America correlated with changes in ocean temperature.......... 459
28. R. H. Fleming: Physical characteristics of the inshore environment .. 483
29. Marston C. Sargent and Theodore J. Walker: Diatom populations associated with eddies off southern California in 1941.......... 490
30. Lionel A. Walford: The case for studying normal patterns in fishery biology.......... 506
31. Oscar E. Sette and Elbert H. Ahlstrom: Estimations of abundance of the eggs of the Pacific pilchard (*Sardinops caerulea*) off southern California during 1940 and 1941.......... 511
32. Henry B. Bigelow and William C. Schroeder: New genera and species of batoid fishes.......... 543
33. Denis L. Fox, Sheldon C. Crane and Bayard H. McConnaughey: A biochemical study of the marine annelid worm, *Thoracophelia mucronata*. Its food, biochromes and carotenoid metabolism.......... 567
34. Wesley R. Coe: Nutrition, environmental conditions, and growth of marine bivalve mollusks.......... 586
35. Claude E. ZoBell and Sydney C. Rittenberg: Sulfate-reducing bacteria in marine sediments.......... 602
36. Wm. D. Urry: Radioactivity of ocean sediments. VII. Rate of deposition of deep-sea sediments.......... 618
37. C. L. Utterback and L. A. Sanderman: Radium analyses of marine sediments in northern Pacific and adjacent waters.......... 635
38. K. O. Emery and J. F. Foster: Water tables in marine beaches.......... 644
39. U. S. Grant: Influence of the water table on beach aggradation and degradation.......... 655
40. Francis P. Shepard: Evidence of world-wide submergence.......... 661
41. Parker D. Trask: Oceanography and geosynclines.......... 679

Errata.......... 686

HARALD ULRIK SVERDRUP—AN APPRECIATION

As a young man Harald Sverdrup spent seven years on Amundsen's MAUD expedition in the icy wastes of the Arctic Sea. The observations he collected during this period form the basis of fifty published papers and constitute much of our factual knowledge about geophysical phenomena in the Arctic. Nevertheless, those years of voluntary exile would be of relatively small interest to his colleagues were it not for their effect on his personal and scientific qualities, and the insight which they give into his character.

The most significant of the MAUD papers, *Dynamics of Tides on the North Siberian Shelf*, was completed for publication during the last winter in the Arctic, without adequate literature and without the benefit of discussion with colleagues. Such an accomplishment was possible only because Sverdrup had developed, to a remarkable degree, internal harmony and simplicity during the long years on the ice. These gave him an objective confidence in his own thinking and the ability to bring his thoughts to well rounded, definite conclusions. Internal harmony between the diverse aspects of a complex personality, between the explorer and the scholar, the naturalist and the theorist, the teacher and the administrator, is still Sverdrup's outstanding characteristic.

Yet the achievement during the expedition of which Sverdrup is most proud is not scientific: after seven years on the MAUD he parted friends with his shipmates. Such a feat must have required all the humor, kindliness, insight and self-discipline which characterize him today. Behind that self-discipline one occasionally glimpses a Norse emotionality, intensity and recklessness.

Sverdrup is impatient with the making of observations unless they are directed toward a definite problem. In all his papers he attempts to wring the last bit of understanding out of each number in a set of data. This frugality must derive partly from the extreme difficulty of taking field observations in the Arctic; partly it reflects his upbringing as a minister's son. Sverdrup's maternal grandfather was a professor of theology, while his paternal grandfather, a minister of the church, had five sons, all of whom became ministers, and three daughters, who married ministers.

Sverdrup was born on November 15, 1888, in Sogndal, Norway, of a family which had long been prominent in the political, religious and scientific life of the country. His cousin Otto Sverdrup was a well known explorer and captain of Nansen's ship FRAM. After a boyhood spent in various parts of western Norway, Harald Sverdrup graduated with honors from the Gymnasium, then spent one year at the Academy of War where he finished as top man in athletics. In 1908 he enrolled at the University of Oslo with the intention of majoring in a subject called "Physical Oceanography and Astronomy,"—with emphasis on the latter. Three years later he was offered a job as assistant to Professor V. Bjerknes, who, with the support of an annual grant from the Carnegie Institution of Washington, was working on the application of hydrodynamics to problems in meteorology and oceanography.

During the next six years Sverdrup published twenty papers, some dealing with dynamic meteorology, others with rather formalistic properties of scalar and vector fields, and one with the radioactivity of sea water in a Norwegian fjord. The outstanding paper of this period is his doctoral dissertation, *Der Nordatlantische Passat*, which gave the first consistent picture of the dynamic and thermodynamic structure of the North Atlantic trade winds. Most of the work on his dissertation was performed while Sverdrup was distractingly hungry. He had accompanied Professor Bjerknes to Leipzig during the war and food was scarce.

At this time Sverdrup agreed to take charge of the scientific work on Roald Amundsen's North Polar Expedition on the MAUD, with additional duties as navigator and part-time cook. He was 29 years old, had spent the last six years on theoretical work, discussing observations taken by others, and was anxious for a change. The expedition left Norway in July 1918, planning to come home in three or four years. Sverdrup did not return to Norway until December 1925.

Upon his return he succeeded V. Bjerknes in the Chair of Meteorology at the Geophysical Institute in Bergen, Norway. In 1931 he became research professor at the Christian Michelsen Institute. Both appointments were made with the understanding that Sverdrup could devote himself to the analysis of his MAUD observations, an activity which occupied the greater part of his time up to 1933.

Among Sverdrup's published papers dealing with the cruise of the MAUD are a general account of the expedition, a popular book on his personal experiences during eight months spent as a lone white man with the nomadic Chukchi tribe of northeastern Siberia, special papers dealing with meteorological observations, tides, currents, physical properties of sea water, sea ice, marine geology, aurora polaris, gravity magnetism, atmospheric electricity, astronomic observations, and *Das Tier- und Vogelleben im Treibeis.* These reports represent a general advance in many fields rather than a far reaching penetration along a narrow scientific salient. The geophysical studies reveal consummate skill in the handling of numerical data, together with a rare degree of ability to recognize physical principles underlying the data and to demonstrate these principles quantitatively by an inelegant and simple yet powerful mathematics. The strength of Sverdrup's later work lies in just this approach: the interpretation and synthesis of numerical measurements. His work is not typically characterized by interpretation of qualitative observations, nor by the development of fundamental theory.

While Sverdrup was engaged in the analysis of the MAUD data, he spent two half-year periods at the Department of Terrestrial Magnetism of the Carnegie Institution of Washington and participated in the analysis of oceanographic data collected on the seventh cruise of the CARNEGIE. In 1928 he married Gudrun Bronn Vaumund. In 1931 he was in charge of the scientific work on the Wilkins Ellsworth North Polar Expedition aboard the submarine NAUTILUS.

Sverdrup spent two months in 1934 with the glaciologist H. W. Ahlmann in the high-lying snow fields of Spitzbergen. This work in glaciology, though brief, resulted in the first quantitative study of the heat budget of glaciers and in several other important investigations. In 1936 he accepted, for a limited period, the Directorship of the Scripps Institution of Oceanography of the University of California, and accordingly took leave of absence from the Michelsen Institute for three years. He did not permanently return to Norway until March 1948.

The twelve years Sverdrup spent in California confronted him with many new problems. He had administrative responsibilities, which he discharged with thoughtfulness and distinction; he gained opportunities to teach, an activity which he enjoys and does exceedingly well. When he first came to California he had one or two students at a time; by the end of the war the need for more trained physical oceanographers had become evident, and Sverdrup organized the first systematic course in oceanography. More than two dozen students who have taken these courses will make his influence felt long after he has left the United States.

Sverdrup's effectiveness as a teacher is due only in part to the clarity and simplicity of his presentation; in larger measure it results from his modesty and his interest in human beings. He makes colleagues of his students rather than subordinates. His cultural interests outside of his work reflect the same humanism—he is concerned primarily with those works of man which give insight into the human heart.

Much of Sverdrup's research in La Jolla was directed toward problems dealing with the air-sea boundary, turbulence in the atmosphere and ocean, and upwelling. He also studied the oceanography of the Pacific, particularly the regions off southern California, and organized an expedition into the Gulf of California. His adaptability to new influences led him to become interested in the inter-relationship between marine biology and physical oceanography, and he collaborated on several papers dealing with marine biology. As an outgrowth of this widened range of interests he became senior author of a comprehensive 1100-page volume, *The Oceans, their Physics, Chemistry, and General Biology;* some of the chapters written by him were rewritten and published in a smaller companion volume, *Oceanography for Meteorologists.* In many ways Chapter XV of *The Oceans,* entitled "The Water Masses and Currents of the Oceans," is his outstanding achievement. In this 157-page chapter Sverdrup's "feeling" for data, his common sense interpretation of physical laws, his experience and remarkable memory are brought to bear upon the problem of dealing with all oceans from a common point of view. In this, as in all his work, he shows a canny judgment in selecting a productive approach to the data and in rejecting the vastly greater number of approaches which at first may seem equally promising but which lead only to inconclusive frustration.

At the outbreak of the war Sverdrup was engaged at the University of California Division of War Research on work dealing with underwater sound. Soon he had to devote his entire energy to the expanding military activities at the Scripps Institution. Among the research problems in which he personally played a leading role were the construction of current charts for life rafts, and the development of methods for predicting sea and swell for amphibious operations.

During his earlier years at Scripps, Sverdrup had been elected a member of the National Academy of Sciences and had received the Agassiz medal. At the end of the war new honors were bestowed upon him. He was elected President of the International Association of Oceanography, Vice-President of the American Geophysical Union, President of the International Commission on Polar Meteorology, and he received the LL.D. degree from the University of California. At the height of his success in America he was invited by the Norwegian Government to assume the Directorship of the new Norwegian Polar Institute. He found it very difficult to make a decision. His acceptance was based, in his own words, on his desire "to assist that small nation . . . during a period when it is struggling to get back on its feet after five years of oppression."

Thirty years ago, as a young man, Harald Sverdrup embarked for the Arctic. Now, at the age of sixty, still young in heart and mind, he is returning to those forbidding Polar regions which have given so many of their secrets to him and have played such a large part in molding his character and his scientific career. The contributors to this volume are but a few of his friends in America who wish him continuing happiness and achievement.

ROGER REVELLE
WALTER MUNK
Scripps Institution of Oceanography
La Jolla, California
June 15, 1948.

BIBLIOGRAPHY OF HARALD ULRIK SVERDRUP

1914 Ausgedehnte Inversionsschichten in der freien Atmosphäre. Veröff. geophys. Inst. Univ. Lpz., (2) *1* (3): 75–100.

—— (Hesselberg, Th., and) Über den Einfluss der Gebirge auf die Luftbewegung längs der Erdoberfläche und auf die Druckverteilung. Veröff. geophys. Inst. Univ. Lpz., (2) *1* (4): 101–116.

—— (Hesselberg, Th., and) Das Beschleunigungsfeld bei einfachen Luftbewegungen. Veröff. geophys. Inst. Univ. Lpz., (2) *1* (5): 117–146.

—— (Hesselberg, Th., and) Die Reibung in der Atmosphäre. Veröff. geophys. Inst. Univ. Lpz., (2) *1* (10): 241–309.

1915 (Hesselberg, Th., and) Beitrag zur Berechnung der Druck- und Massenverteilung im Meere. Bergens Mus. Aarb. (1914–15), No. 14: 1–18.

—— (Hesselberg, Th., and) Die Stabilitätsverhältnisse des Seewassers bei vertikalen Verschiebungen. Bergens Mus. Aarb. (1914–15), No. 15: 1–16.

—— (Hesselberg, Th., and) Die Windänderung mit der Höhe vom Erdboden bis etwa 3000 m Höhe. Beitr. Phys. frei. Atmos., Bd. 7, Heft 3: 156–166.

1916 Druckgradient, Wind und Reibung an der Erdoberfläche. Ann. Hydrogr., Berl., *44:* 413–427.

—— Stationäre Bewegungsfelder. Met. Z., *33*, Heft 5: 208–210.

—— Der feucht-adiabatische Temperaturgradient. Met. Z., *33*, Heft 6: 265–272.

—— Über Mittelwerte von Vektorpaaren mit Anwendungen auf meteorologische Aufgaben. Met. Z., *33*, Heft 9: 411–420.

1917 Der nordatlantische Passat. Veröff. geophys. Inst. Univ. Lpz., (2) *2* (1): 1–96, 19 tables.

—— (and J. Holtsmark) Über die Reibung an der Erdoberfläche und die direkte Vorausberechnung des Windes mit Hilfe der hydrodynamischen Bewegungsgleichungen. Veröff. geophys. Inst. Univ. Lpz., (2) *2* (2): 97–141.

—— (and J. Holtsmark) Über die Beziehung zwischen Beschleunigungen und Gradientenänderungen und ihre prognostische Verwendung. Veröff. geophys. Inst. Univ. Lpz., (2) *2* (3): 143–171.

—— Über die Korrelation zwischen Vektoren mit Anwendungen auf meteorologische Aufgaben. Met. Z., *34*, Heft 8/9: 285–291.

—— Zur Bedeutung der Isallobarenkarten. Ann. Hydrogr., Berl., *45:* 325–329.

1918 Der nordatlantische Passat (W. Köppen, nach HUS). Ann. Hydrogr., Berl., *46:* 31–36.

—— Die Beziehung der elfjährigen Klimaschwankungen zur Sonnentätigkeit. Ann. Hydrogr., Berl., *46:* 191–193.

—— Einige Untersuchungen über die Radioaktivität des Seewassers in den Fjorden in der Nähe von Bergen (Norwegen). Bergens Mus. Aarb. (1917–18), No. 12: 1–5.

—— Über den Energieverbrauch der Atmosphäre. Veröff. geophys. Inst. Univ. Lpz., (2) *2* (4): 173–196.

1921 Blandt rentsjuktsjere og lamuter (pp. 257–391); In Amundsen: Nordost-Passagen. Gyldendalske Boghandel, Kristiania. 467 pp.

1922 Customs of the Chukchi natives of northeastern Siberia. J. Wash. Acad. Sci., *12:* 208–212.

—— Maud-ekspeditionens videnskabelige arbeide 1918–19 og nogen av dets resultater. Naturen, 46de aarg., No. 1/2: 5–32.

—— Maud-ekspeditionens videnskabelige arbeide 1918–19 og nogen av dets resultater. Naturen, 46de aarg., No. 3/4: 65–88.

—— Meteorology on Captain Amundsen's present Arctic Expedition. Mon. Weath. Rev. Wash., *50:* 74–75.

—— (and C. R. Duvall) Results of magnetic observations on the "Maud Expedition," 1918–1921. Terr. Magn. atmos. Elect., *27:* 35–56.

1925 The North-polar cover of cold air. Mon. Weath. Rev. Wash., *53:* 471–475.

1926 Maud-ekspeditionen 1918–1925. Ymer (Tidskr. Svenska Sällsk. Antropologi Geografi), *46*, Heft 1: 1–18.

—— "Maud"-ekspeditionens videnskabelige arbeide 1922–1925. Naturen, *50:* 161–180.

—— Scientific work of the "Maud" expedition, 1922–1925. Sci. Month., N. Y., *22:* 400–410. (Also in, Rep. Smithson. Inst. [1926]; pp. 219–233, 1927.)

—— (and O. Dahl) Two oceanographic current-recorders designed and used on the "Maud" expedition. J. opt. Soc. Amer. and Rev. Sci. Instr., *12:* 537–545.

—— The tides on the North Siberian shelf: Their bearing on the existence of land in the Arctic Sea, and their dynamics. J. Wash. Acad. Sci., *16:* 529–540.

—— Tre ar i isen med "Maud." Glydendal Norsk Forlag, Oslo. 285 pp.

1927 Ergebnisse der Messungen des Potentialgefälles auf der "Maud"-Expedition. Z. Geophys., Jahrg. 3, Heft 2/3: 93–102.

—— Dynamic of tides on the North Siberian shelf. Geofys. Publ., *4* (5): 1–75.

—— Magnetic, atmospheric-electric, and auroral results, Maud expedition, 1918–1925. Dept. Terr. Magn., Publ. Carneg. Inst. Wash., VI, 175: 309–524. (Included in Sci. Res. Norwegian . . . "Maud" 1918–1925, as vol. 1b).

—— (Wait, G. R., and) Preliminary note on electromotive forces possibly produced by the earth's rotating magnetic field and on observed diurnal-variation of the atmospheric potential-gradient. Terr. Magn. atmos. Elect., *32:* 73–83.

—— Nordenskiölds hav og det Øst-Sibiriske hav. Norsk geogr. Tidsskr., *1*, Heft 6/7: 321–335.

—— Scientific work of the "Maud" expedition, 1922–1925. Rep. Smithson. Inst. (1926), pp. 219–233.

1928 Finn Malmgrens videnskabelige virke. Ymer, *48*, Heft 3: 246–252.

—— Aufgaben, Bemannung und Ausrüstung einer wissenschaftlichen Beobachtungsstation auf dem Treibeis bei 1–2 jähriger Überwinterung in der inneren Arktis. Arktis, Jahrg. 1928, Heft 1/2: 29–36.

—— Die Eistrift im Weddelmeer. Ann. Hydrogr., Berl., *56:* 265–274.

—— On the importance of auroral photographs taken from one station. Terr. Magn. atmos. Elect., *33:* 198–202.

—— Die Renntier-Tschuktschen. Mitt. geogr. Ges. Hamburg, Bd. 39: 97–135.

—— Minnetale over Roald Amundsen. Arb. norske VidenskAkad., pp. 125–129.

—— Results of astronomical observations. Sci. Res. Norweg. N. Pol. Exped. "Maud," *1* (3): 24 pp.

—— The wind-drift of the ice on the North-Siberian shelf. Sci. Res. Norweg. N. Pol. Exped. "Maud," *4* (1): 46 pp.

1929. Currents on the North Siberian shelf. Beretn. 18th skand. Naturforskerm., pp. 522–527.

—— Polferden med "Graf Zeppelin." Naturen, *53:* 353–367.

—— (Hesselberg, Th., and) Über die Genauigkeit der Berechnung der Druck-und Massenverteilung und der Stabilitätsverhältnisse im Meere. Ann. Hydrogr., Berl., *57:* 73–75.

—— The waters on the North-Siberian shelf. Sci. Res. Norweg. N. Pol. Exped. "Maud," *4* (2): 131 + 75 pp.

1930 The bottom water on the North-Siberian shelf. Congr. Internac. Oceanog., Hidrogr. mar. e Hidrol. cont., Sevilla, 1–7 Mayo, 1929, pp. 331–336d.

—— Dyrelivet i driv-isen. Naturen, *54:* 133–145.

—— Fridtjof Nansen. Arktis, Jahrg. 1930, Heft 1/2: 1–4.

—— Meteorology; Tables. Sci. Res. Norweg. N. Pol. Exped. "Maud," *3* (2): 527 pp.

—— Some aspects of oceanography. Sci. Mon. N. Y., *31:* 19–34.

—— Some oceanographic results of the "Carnegie's" work in the Pacific—The Peruvian current. Trans. Amer. geophys. Un., pp. 257–264. (Reprint. in, Hydrogr. Rev., *8:* 240–244, 1931).

1931 Audibility of the Aurora polaris. Nature, London, *128:* 457.

—— Fridtjof Nansen som videnskapsmann. Norsk geogr. Tidsskr., *3:* 306–313.

—— Diurnal variation of temperature at polar stations in the spring. (Gerlands) Beitr. Geophys., *32:* 1–14 (Köppen-Band I).

—— Internationale Studiengesellschaft zur Erforschung der Arktis mit dem Luftschiff (Aeroarctic). II. Die meteorologischen Untersuchungen und Ergebnisse der "Maud"-expedition. Petermanns Mitt., Ergänzungs. 191: 63–68.

—— The deep-water of the Pacific according to the observations of the Carnegie. Rept. to Sect. Oceanogr., Intern. Geod. Geophys. Un., Stockholm; Dept. terr. Magn., Carnegie Inst. Wash., pp. 87–94.

—— The origin of the deep-water of the Pacific Ocean as indicated by the oceanographic work of the Carnegie. (Garlands) Beitr. Geophys., *29:* 95–105.

—— The origin of the deep-water of the Pacific Ocean as indicated by the oceanographic work of the *Carnegie*. Rept. to Sect. Oceanogr., Intern. Geod. Geophys. Un., Stockholm; Dept. Terr. Magn., Carnegie Inst. Wash., pp. 101–109.

—— Resultater av Maudferdens oseanografiske undersøkelser. Naturen, *55:* 1–16.

—— Scientific results of the Andrée-Expedition. I. Drift-ice and ice-drift. Geogr. Ann., Stockh., *13:* 121–140.

—— Snedekkets termiske egenskaper. Beretn. Michelsens Inst. Vidensk., Bd. 1 (3): 1–21.

—— Das Tier- und Vogelleben im Treibeis. Petermanns Mitt., Heft 1/2: 13–16.

—— Die wissenschaftlichen Arbeiten auf der Wilkins-Ellsworth Expedition 1931. Arktis, Jahrg. 1931, Heft 4: 49–50.

—— Hvorledes og hvorfor med "Nautilus." Gyldendal Norsk Forlag., Oslo. 183 pp.

1932 Als Meeresforscher mit dem Unterseeboot "Nautilus" im Nordpolargebiet. Das Meer, Bd. I, Polarbuch. E. S. Mittler & Son, Berlin. pp. 1–22.

—— Arbeider i luft- og havforskning. Beretn. Michelsens Inst. Vidensk., Bd. 2 (5): 1–20.

—— Wärmehaushalt und Austauschgrösse auf Grund der Beobachtungen der "Maud"-expedition. Beitr. Phys. frei. Atmos. (Bjerknes Festschrift), *19:* 276–290.

1933 General report of the expedition. Sci. Res. Norweg. N. Pol. Exped. "Maud," *1* (1): 22 pp.

—— Pendulum observations near Cape Chelyuskin. Sci. Res. Norweg. N. Pol. Exped. "Maud," *1a* (8): 9 pp.

—— Meteorology; Discussion. Sci. Res. Norweg. N. Pol. Exped. "Maud," *2* (1): 331 pp.

—— Geofysiske undersökelser, saerlig over vindens betydning for havströmmene. Beretn. Michelsens Inst. Vidensk., Bd. 3 (5): 1–25.

—— Scientific results of the "Nautilus" expedition, 1931. I. Introduction and narrative; II. Oceanography. Pap. phys. Oceanogr. Meteorol., *2* (1): 1–63.

—— Naturvidenskap og religion. Fritt Ord., *3:* 107–111.

—— On vertical circulation in the ocean due to the action of the wind with application to conditions within the Antarctic circumpolar current. 'Discovery' Rep., *7:* 139–170.

—— Vereinfachtes Verfahren zur Berechnung der Druck- und Massenverteilung im Meere. Geofys. Publ., *10* (1): 1–9.

1934 The circulation of the Pacific. Proc. 5th Pacif. Sci. Congr., Canada (1933), pp. 2141–2145.

—— Oversikt over "Maud"-ekspedisjonens videnskapelige resultater. Beretn. Michelsens Inst. Vidensk., IV (7): 1–18 (Also in Naturen, *58:* 1–15, 1934).

—— Videnskapens Bakgrunn. Studentersamfundet it Trondhjem Smaskrifter, Nr. 2: 15 pp.

—— Wie entsteht die Antarktische Konvergenz? Ann. Hydrogr. Berl., *62:* 315–317.

1935 The temperature of the firn on Isachsen's Plateau, and general conclusions regarding the temperature of the glaciers on West-Spitsbergen. Part III of Scientific results of the Norwegian-Swedish Spitsbergen Expedition in 1934. Geogr. Ann., Stockh., *17*, Heft 1/2: 53–88.

—— The ablation on Isachsen's Plateau, and on the Fourteenth of July Glacier in relation to radiation and meteorological conditions. Part IV of Scientific results of the Norwegian-Swedish Spitsbergen Expedition in 1934. Geogr. Ann., Stockh., *17*, Heft 3/4: 145–166.

—— Temperaturen i Vest-Spitsbergens breer. Naturen, *59:* 239–248.

—— Übersicht über das Klima des Polarmeeres und des Kanadischen Archipels. In, Handbuch der Klimatologie herausgegeben von W. Köppen und R. Geiger. Bd. II, Teil K: 1–30.

—— Varmeutvekslingen mellem en sneflate og luften. Beretn. Michelsens Inst. Vidensk., Bd. 5 (3): 1–33.

—— Zum Wärmehaushalt der Gletscher auf West-Spitzbergen. Met. Z., *52* (12): 495.

1936 Austausch und Stabilität in der untersten Luftschicht. Met. Z., *53*, Heft 1: 10–15.

—— The eddy conductivity of the air over a smooth snow field. Geofys. Publ., *11* (7): 1–69.

—— Das maritime Verdunstungsproblem. Ann. Hydrogr. Berl., *64:* 41–47.

—— Results of the meteorological observations on Isachsen's Plateau. Part IV of Scientific results of the Norwegian-Swedish Spitsbergen Expedition in 1934. Geogr. Ann., Stockh., *18*, Heft 1/2: 34–47.

—— Turbulensforskning i laboratoriet og i naturen. Beretn. Michelsens Inst. Vidensk., VI (5): 1–21.

1937 Oceanographic research at the Scripps Institution of Oceanography during April 1936 to April 1937. Trans. Amer. geophys. Un., pp. 210–216.

—— On the evaporation from the oceans. Science, *85* (2210): 439 (abstr.). (Also: Bull. Amer. met. Soc., *18:* 400–401).

—— On the evaporation from the oceans. J. Mar. Res., *1:* 3–14.

—— (D. L. Fox, H. U. S., and J. P. Cunningham) Rate of water propulsion by the California mussel. Biol. Bull. Woods Hole, *72:* 417–438.

—— The work at the Scripps Institution of Oceanography. Collecting Net, *12:* 57, 60, 61.

1938 Notes on erosion by drifting snow and transport of solid material by sea ice. Amer. J. Sci., (5) *35:* 370–373.

—— On the explanation of the oxygen minima and maxima in the oceans. J. Cons. int. Explor. Mer, *13:* 163–172. (Review: Zbl. Geophys., Bd. 3, Heft 4: 161.)

—— On the process of upwelling. J. Mar. Res., *1:* 155–164.

—— Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1937 to April 1938. Trans. Amer. geophys. Un., *19:* 238–242.

—— Oceanographic problems off the coast of California. Trans. Amer. geophys. Un., *19:* 173–174.

—— Hos Tundra-Folket. Gyldendal Norsk Forlag, Oslo. 175 pp.

1939 Oceanic circulation. Proc. 5th int. Congr. appl. Mech. (1938), pp. 279–293.

—— On the influence of stability and instability on the wind profile and the eddy conductivity near the ground. Proc. 5th int. Congr. appl. Mech. (1938), pp. 369–372.

—— Physics and geophysics: With special reference to problems in physical oceanography. Univ. California at Los Angeles, Faculty Research Lecture, March 23, 1938. Univ. Calif. Press. 23 pp.

—— Second note on the logarithmic law of wind structure near the ground. Quart. J. R. met. Soc., *65:* 57–60.

—— Response of the medallist. (Presentation of the Agassiz Medal to H. U. Sverdrup.) Science, *90* (2324): 24–27.

—— Cruises of the E. W. Scripps in 1939. Sci. Mon., N. Y., *49:* 389–391.

—— Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1938 to April 1939. Trans. Amer. geophys. Un., *20:* 422–427.

—— (and W. E. Allen) Distribution of diatoms in relation to the character of water masses and currents off southern California in 1938. J. Mar. Res., *2:* 131–144.

—— Lateral mixing in the deep water of the South Atlantic Ocean. J. Mar. Res., *2:* 195–207.

1940 Hydrology; Discussion. Rep. Brit. Austr. N. Zealand Antarc. Exped., 1921–1931, (A) *3*, Pt. 2, Sect. II: 87–126.

—— The currents of the Pacific Ocean and their bearing on the climates of the coasts. Science, *91* (2360): 273–282.

—— General remarks on turbulence in the atmosphere and the ocean. P. V. Ass. Océanogr. Phys., *3:* 118.

—— The Gulf of California. P. V. Ass. Océanogr. Phys., *3:* 170–171.

—— Do permanent deep-sea currents exist? P. V. Ass. Océanogr. Phys., *3:* 182–183.

—— The Arctic regions. Publ. sci. Ass. Océanogr. phys., *8:* 50–53.

—— On the annual and diurnal variation of the evaporation from the oceans. J. Mar. Res., *3:* 93–104.

—— Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1939 to April 1940. Trans. Amer. geophys. Un., *21:* 343–346.

—— The unity of the sciences of the sea. Sigma Xi Quart., *28* (3): 105–115.

—— Trail blazing in the Pacific (Research in oceanography). Calif. Month., *45:* 10, 11, 28.

—— Activities of the Scripps Institution of Oceanography, La Jolla, California. Proc. 6th Pacif. Sci. Congr., Calif., 1939, *3:* 114–123.

—— The Gulf of California. Preliminary discussion of the cruise of the E. W. Scripps in February and March, 1939. Proc. 6th Pacif. Sci. Congr., Calif., 1939, *3:* 161–166.

1941 Water masses and currents of the North Pacific Ocean. Science, *93* (2419): 436.

—— The influence of bottom topography on ocean currents (pp. 66–75); In, Contributions to Applied Mechanics and Related Subjects. Theodore von Kármán Anniversary Volume. California Inst. Technology, Pasadena. 337 pp.

—— The Pacific Ocean. Science, *94* (2439): 287–293.

—— An analysis of the ocean currents off the American west coast between 40° N and 40° S (pp. 17–20). In, Proceedings of the Dedicatory Exercises, Hancock Hall. Univ. Chron. Ser., No. 7. Univ. S. Calif. Press, Los Angeles. 37 pp.

—— (and R. H. Fleming) The waters off the coast of southern California March to July, 1937. Bull. Scripps Inst. Oceanogr., *4* (10): 261–378.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1940 to April 1941. Trans. Amer. geophys. Un., *22:* 490–494.

1942 (and Staff) Oceanographic observations on the "E. W. Scripps" cruises of 1938. Rec. Observ. Scripps Inst. Oceanogr., *1* (1): 1–63.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1941 to April 1942. Trans. Amer. geophys. Un., *23* (2): 323–325.

—— Oceanography for meteorologists. Prentice-Hall, Inc., New York. 246 pp.

—— (M. W. Johnson, and R. H. Fleming) The Oceans: Their physics, chemistry, and general biology. Prentice-Hall, Inc., New York. 1087 pp.

1943 (and Staff) Oceanographic observations of the Scripps Institution in 1939. Rec. Observ. Scripps Inst. Oceanogr., *1* (2): 65–159.

—— On the ratio between heat conduction from the sea surface and heat used for evaporation. Ann. N. Y. Acad. Sci., *44:* 81–88.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1942 to April 1943. Trans. Amer. geophys. Un., *24:* 244–246.

1944 (and Staff) Oceanographic observations of the "E. W. Scripps" cruises of 1940. Rec. Observ. Scripps Inst. Oceanogr., *1* (3): 161–248.

—— The California Current (pp. 97–111); In, Science in the University. Univ. Calif. Press, Berkeley and Los Angeles. 332 pp.

—— (*et al.*) Observations and results in physical oceanography. Sci. Res. Cruise VII "Carnegie," 1928–1929. Oceanogr., I-A: 1–156.

1945 (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1943 to April 1944. Trans. Amer. geophys. Un., *26:* 605.

—— Oceanography (pp. 1029–1056); In, Handbook of Meteorology. McGraw-Hill Book Co., Inc., New York. 1068 pp.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1944 to April 1945. Trans. Amer. geophys. Un., *26:* 127–128.

1946 The humidity gradient over the sea surface. J. Meteorol., *3:* 1–8.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1945 to April 1946. Trans. Amer. geophys. Un., *27:* 571–573.

—— (and W. H. Munk) Empirical and theoretical relations between wind, sea, and swell. Trans. Amer. geophys. Un., *27:* 823–827.

—— (and W. H. Munk) Theoretical and empirical relations in forecasting breakers and surf. Trans. Amer. geophys. Un., *27:* 828–836.

1947 (and W. H. Munk) Wind, sea, and swell: Theory of relations for forecasting. Tech. Rep. U. S. Hydrogr. Off., No. 1; H. O. Publ., No. 601: 44 pp.

—— New international aspects of oceanography. Proc. Amer. phil. Soc., *91:* 75–78.

—— Note on the correction of reversing thermometers. J. Mar. Res., *6:* 136–138.

—— Period increase of ocean swell. Trans. Amer. geophys. Un., *28:* 407–417.

—— Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the eastern Pacific. Proc. nat. Acad. Sci., *33:* 318–326.

—— (and Staff) Research within physical oceanography and submarine geology at the Scripps Institution of Oceanography during April 1946 to April 1947. Trans. Amer. geophys. Un., *28:* 801–802.

—— (and R. H. Fleming) Atlantic Ocean. Encyclopaedia Britannica, 5 pp.

—— (and R. H. Fleming) Indian Ocean. Encyclopaedia Britannica, 2 pp.

—— (and Staff) Oceanographic observations on the E. W. Scripps cruises of 1941. Rec. Observ. Scripps Inst. Oceanogr., *1* (4): 249–408.

—— Wind, sea, and swell. Proc. roy. Canad. Inst. (Sess. 1946–1947), (3A) *12:* 55–56.

In press. Physical oceanography of the Polar Sea. In, North American Arctic.

—— (M. W. Johnson and R. H. Fleming) Ocean and oceanography. Encyclopaedia Britannica.

—— (and R. H. Fleming) Pacific Ocean. Encyclopaedia Britannica.

SEARS FOUNDATION FOR MARINE RESEARCH
BINGHAM OCEANOGRAPHIC LABORATORY, YALE UNIVERSITY

JOURNAL OF
MARINE RESEARCH

VOLUME VII 1948 NUMBER 3

A POSSE AD ESSE

By

DANIEL MERRIMAN[1]

Bingham Oceanographic Laboratory
Yale University
New Haven, Connecticut

"From possibility to reality" seems an appropriate phrase to describe the development of oceanography in the past three-quarters of a century. It is of little more than academic interest to pick a particular date or event as marking the birth of a field of endeavor; however, for oceanography, the late 1860's and the early 1870's unquestionably mark the period which gave the greatest impetus to research on the chemistry, physics and biology of the sea. Seventy-five years ago Her Majesty's Ship CHALLENGER was well started on her round-the-world expedition which covered 69,000 miles and took three and a half years. Seventy-five years ago, also, there was published "The Depths of the Sea" by Charles Wyville Thomson (1). Oceanography at that time was a possibility. That it is now a reality is self-evident—although too often taken for granted, with little realization of the relatively short space of time it took to accomplish the transition "a posse ad esse." If Thomson's book ". . . may be regarded as the first general textbook on oceanography" (2, p. 41), the comprehensive volume by Sverdrup, Johnson and Fleming (3) would appear to be the last. This field is now so large that it is doubtful if another similar account will ever again be placed between the two covers of a single book. It is appropriate in honoring Dr. Sverdrup that we should look back seventy-five years to the bare beginnings of the field to which he has made such notable contribution.

---

[1] It is with infinite pleasure that I acknowledge my indebtedness to my wife, Mary Wieland Merriman, whose painstaking searches for original material made this work possible.

One of the major problems in the third quarter of the 19th century was the question as to the depths at which animal life might exist in the ocean. The position is summarized nicely in a review (4) of Thomson (1).

> In no department of zoology has there been of late years a greater revolution than in that to which belongs the existence of life at great ocean depths. Not twenty years ago, it may be said, with almost all naturalists as with one not less distinguished than the late Edward Forbes, the belief in a zero of life, or a depth below which no animal existence was possible, was an article of the scientific creed. Under conditions of pressure, of defective light, temperature, and aëration, such as were held to prevail in the abysses of the deep sea, nothing approaching to known varieties of animal life could, it was thought, by possibility exist. Cases, however well authenticated, of animals comparatively high in the scale of life having been brought up by sounding lines from great depths, were set aside even by men of eminent standing in science, as having been due to careless observation, or to the fact of the creatures having got entangled in the line while floating near the surface. It is to the operations involved in laying and subsequently fishing up the Atlantic telegraph cable that the first great impulse towards an opposite belief is to be traced. Solitary or infrequent cases had indeed occurred to stir the thoughts of naturalists. There was the beautiful Medusa's head, fished up by Sir John Ross from one thousand fathoms in Baffin's Bay, lat. 73° 37′ N., September 1, 1818; and there were the rich hauls of coralines, and other varieties of invertebrate life, obtained by Sir J. C. Ross in equally far Southern latitudes and from much the same depths, suggesting to those able explorers, though contrary to the received opinion of naturalists, that the lowest depth to which we might succeed in penetrating would be found to teem with animal life. The minute samples of Atlantic ooze brought up by the sounding apparatus introduced by Mr. Brooke in 1854 made known the existence and diffusion of microscopic forms not distinguishable from the fossil constituents of chalk, which the discovery of identical specimens from the Pacific showed to be in all probability distributed over the whole ocean floor. Doubt still existed whether these organisms had their life at the bottom of the sea, or were precipitated in fine showers after death. The opinion of Ehrenberg in favour of the first of these hypotheses was balanced by that of Professor Baily, of West Point, on the contrary side; Professor Huxley in 1858, in his report upon the problem to the Admiralty, expressing himself guardedly as on the whole inclined to believe that the Globigerinae and other foraminifers and diatoms really lived at these depths. Dr. Wallich's dredgings in the North Atlantic, though not conclusive of the question, contributed much to strengthen the same belief, which was generally held to be established for good when Mr. Fleeming Jenkin showed the manifold specimens brought up with the submerged cable between Bona and Sardina from depths extending to 1,200 fathoms, among which Professor Allman readily identified fifteen animal forms, including the ova of a cephalopod. Subsequent examination of organisms derived from the same source enabled M. Milne Edwards to pronounce the problem finally solved.

In the spring of 1868 Charles Wyville Thomson[2] was working in Ireland with W. B. Carpenter, then one of the Vice Presidents of the Royal Society, on the structure and development of crinoids. Together they decided to urge upon the Admiralty the importance of providing ". . . a vessel properly fitted with dredging gear and all necessary scientific apparatus, that many heavy questions as to the state of things in the depths of the ocean which were still in a state of uncertainty, might be definitely settled" (1, p. 50).

As a result of Carpenter's and Thomson's proposition to the Admiralty, Her Majesty's steam vessel LIGHTNING was assigned for the purpose of carrying out dredging operations in July 1868. The oft-quoted description of the vessel (1, p. 57) merits repetition here:

> The surveying ship 'Lightning' was assigned for the service—a cranky little vessel enough, one which had the somewhat doubtful title to respect of being perhaps the very oldest paddle-steamer in her Majesty's navy. We had not good times in the 'Lightning.' She kept out the water imperfectly, and as we had deplorable weather during nearly the whole of the six weeks we were afloat, we were in considerable discomfort. The vessel, in fact, was scarcely seaworthy, the iron hook and screw-jack fastenings of the rigging were worn with age, and many of them were carried away, and on two occasions the ship ran some risk. Still the voyage was on the whole almost pleasant.

In the six weeks at their disposal, just ten days were decent enough for dredging in the open sea and on only four of these were they in depths over 500 fathoms. Nonetheless, they revealed varied, abundant, and hitherto unknown animal life down to 650 fathoms, as well as deep-water masses of different temperatures; this was sufficient to convince the Admiralty that another vessel should be assigned for the same purpose in 1869. The PORCUPINE was more seaworthy, and she made several cruises in the North Atlantic and Mediterranean during the summers of 1869 and 1870 under the direction of Thomson, Carpenter, and Gwyn Jeffreys. Dredging first at 1,476 and later at 2,435 fathoms, the existence of animal life at these depths was proven beyond question. "In the Mediterranean Dr. Carpenter found the conditions of temperature and of the distribution of animal life entirely exceptional, as might have been to a certain extent anticipated from the exceptional circumstances of that land-locked sea" (1, p. 179).

---

[2] Thomson (5), born March 5, 1830, had received his education at the University of Edinburgh from 1845 to 1850, where his chief interests were botany, geology and zoology. Although he took no degree, his standing was such that he immediately became Lecturer on Botany in King's College at the University of Aberdeen. In 1853 he was made Professor of Natural History in Queen's College, Cork, and in 1854 was transferred to the Chair of Mineralogy and Geology at Belfast, where in 1860 he became Professor of Zoology and Botany. Such, in brief, was the background of the man who was destined to play a leading role in the development of oceanography from 1868 until his death in 1882.

In 1870 Thomson became Professor of Natural History at the University of Edinburgh, thus following in the footsteps of Edward Forbes who had named the region below 300 fathoms the "Azoic zone." Some modern writers who contradict the theories of their predecessors would do well to read Thomson on Forbes (1, pp. 6–18) as a model of expression and fairness.

There followed the arrangements for the CHALLENGER Expedition, culminating in the appointment of Thomson as director of the civilian scientific staff, and the departure of Her Majesty's Ship in December 1872. The details of the ship, her voyage and discoveries are well known to us at first hand through the works of several authors (6, 7). However, the degree of public interest and the attention these first truly oceanographic investigations drew at the time deserve mention. The first paragraph of a column-and-a-half article (8) in the The (London) Times just before the CHALLENGER sailed is quoted herewith:

> An early day of the ensuing week will witness the departure of the most important naval surveying Expedition which has ever been sent forth by any country. We have from time to time recorded the results obtained by Dr. Carpenter, Professor Wyville Thomson, and Mr. Gwyn Jeffreys, in the brief voyages of the Lightning and the Porcupine; results which showed previously-unsuspected variations in the deep sea temperature, the existence of a general oceanic circulation, the presence of life at the greatest depths, and the active progress of submarine chalk formated even in the present day. The great scientific and practical importance of the facts revealed by these short and imperfect inquiries was such as to render their continuance a matter of national concern, and Dr. Carpenter, who had from the first been most active in their prosecution, exerted himself to bring the whole subject under the consideration of the Government. Mr. Lowe, however strict in protecting the public purse against demands for which private enterprise may be legitimately called upon, fully recognizes that matters which it is fitting for the country to undertake should be carried through with every advantage that money can secure, and he gave early intimation of his approval of the proposed Expedition. The Admiralty, in like manner, asked only the official approval of the Royal Society; and, as soon as this was given, Mr. Goschen lent himself heartly to the scheme. The suggestion made to the Royal Society was that a ship should be fitted out for an expedition of three or four years' duration, during which soundings, themometric (*sic*) observations, dredging, and chymical examination of sea water, should be carried on continuously, with a view to a more perfect knowledge of the physical and biological conditions of the great ocean basins, and in order to ascertain their depth, temperature, specific gravity, and chymical character. At the same time it was recommended that observations should be made on the direction and velocity of the great drifts and currents, especially those of the Gulf Stream, the Equatorial and Japan, both at the surface and in intermediate strata, as well as on the fauna of the

> deep water, and on the zoology and botany of those portions of the globe which are at present comparatively unknown.

For the next three and a half years the Times carried frequent and detailed articles on the progress of the CHALLENGER, her itinerary and accomplishments, and various incidents of the voyage.

Further indication of the general interest in the expedition is to be found in Punch, or the London Charivari; Fig. 1 is a reproduction of the sketch accompanying the Preface to Volume 63 (9), and the Preface itself is the somewhat ribald conversation between Mr. Punch and the Captain on the occasion of the sailing of the CHALLENGER.

Early in 1873, several months after the CHALLENGER had sailed, Thomson's "The Depths of the Sea" was published. The extensive reviews of the book in contemporary papers and journals again provide evidence of the widespread interest in oceanographic research. All of these reviews give high praise to Thomson's researches and powers of expression. Two typical quotations follow:

> . . . we already know a good deal more of the sea—its depths, currents, temperature, density, and animal and vegetable life—than was known only a few years since; and if all goes well with the expedition of the *Challenger*, which recently left our shores, a knowledge of this department of science will

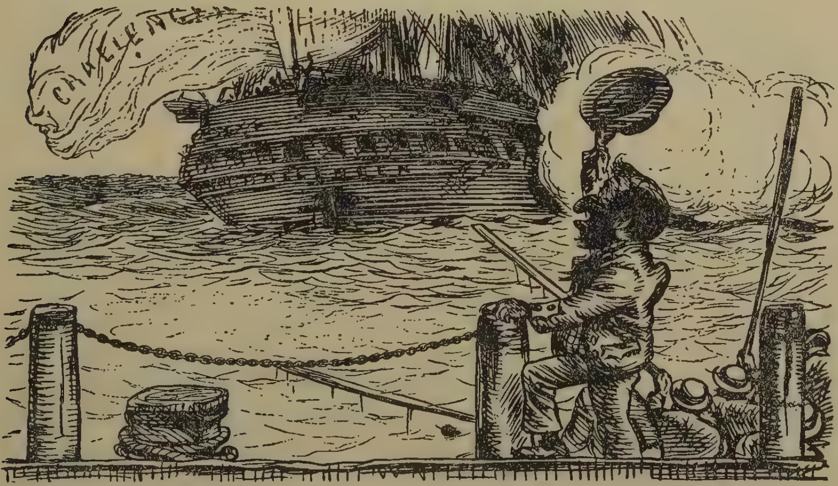


Figure 1. " 'Good luck to you all!' shouted MR. PUNCH, for the last time, and he shouted in the name of all HER MAJESTY'S subjects." (9, p. IV). Reproduced by permission of the Proprietors of PUNCH.

> ere long be materially extended. As regards the facts lately acquired, those who happen to be curious on a subject of so much interest should peruse the work lately published, *Depths of the Sea*, by Dr. C. Wyville Thomson, a goodly octavo, embellished with wood-engravings, and abounding in what we would call very pleasant reading. As among the many thousands we address, few may have an opportunity of seeing the book, we shall endeavour to give some idea of its character and purport. . . .
>
> We unfortunately have not space to pursue this interesting subject, which is treated in its various details in the work before us. The book is an important contribution to physical science, and we cordially recommend it for attentive perusal (10).

> It is a rare treat to the reviewer to be able to introduce to the public of this age a book which really enlarges the boundaries of our knowledge, and is the result of original investigation conducted with the care, caution, and independence which are the elements which compose the atmosphere of true science. Overwhelmed as we are with publications of the parasitical order, which coil around the stems of classic literature and run to the limit of endurance, it is refreshing to find that something both new and true, and which can be startling without losing its simplicity, can still be told, and told so well, by one who has just started with fairer and fuller appliances to reap fresh harvests in the same almost untouched, and almost exhaustless field.
>
> The information here given certainly marks an epoch in the chronicles of science. It establishes by experiment data which completely reverse not only popular notions, but also the sober instructions which the science of yesterday unhesitatingly inculcated. It shows also how the resources of a Government are both necessary and efficient for the furtherance of science. If that cranky little surveying ship the *Lightning*, and afterwards the better equipped but still small and defective vessel, the *Porcupine*, could do so much, what may we not expect from the naval equipment of a nation which, while it boasts of being the mistress of the seas, has allowed its vessels and their crews to lounge about the world like the idlers of a village green, when they might have been converted into trained explorers, without at all impairing their combative qualities (11).

The reviews of Thomson's book are also full of interesting side lights, which reflect the times—a bare 14 years after Darwin's "Origin of Species," and which include many significant and amusing comments. Typical samples follow:

> When one reads Prof. Thomson's account of dredging apparatus and engines, and calls to mind the elaborate arrangements which have been made on board the Challenger, it is interesting to compare the following passage from the narrative of the voyage of the Rattlesnake, which was sent on surveying work to the North Australian coast in 1846, and had on board as naturalist Mr. Macgillivray,—who wrote the history of the voyage,—as surgeon, no less distinguished a naturalist than Mr. Huxley. At Rio de Janeiro they determined to try their luck:—
>
> "I had looked forward with eager anticipation to the result of the first dredging of the voyage. *None of the ship's boats could be spared*, so I hired one pulled by four negro slaves, who, although strong active fellows, had great objections to straining their backs at the oar when the dredge was down.

No sieve having been supplied, we were obliged to sift the contents of the dredge through our hands,—a tedious and superficial mode of examination. Still some fine specimens of a curious flat sea-urchin (*Encope marginata*), and a few shells, encouraged us to persevere. Two days after, Mr. Huxley and myself set to work in Botafogo Bay, provided with a *wire-gauze meat-cover*, and a curious machine for cleaning rice; these answered capitally as substitutes for sieves, and enabled us, by a thorough examination of the contents of the dredge, to detect about forty-five species of mollusca and radiata, some of which were new to science."

The Lords of the Admiralty have, we are glad to know, learned since that time to render it unnecessary for naturalists sent out by them to borrow meat-covers and machines for cleaning rice (12).

One of the most singular things about this deep-sea fauna is that it is by no means universally destitute of eyes. A mollusc, dredged up from 2,000 fathoms, "had a pair of well-developed eyes on short footstalks." Where there are eyes, there must one would suppose—equally on Darwinian or on teleological principles—be light. But it is scarcely possible that sun-light can visit these abysses, and Mr. W. Thomson is accordingly inclined to believe that whatever light affords an exercise to these deeply buried organs of sight may be due only to phosphorescence (13).

With nearly all the leading naturalists of his time, Forbes was a believer in the immutability of species. But upon this point the author alludes to the great change of opinion which has taken place within the last ten or twelve years, due to the ability and candor with which the question has been treated by Darwin and Wallace, and to the genius of Professor Haeckel, Dr. Fritz Müller, and others of their enthusiastic disciples and commentators. The author, however, states the case too strongly in asserting "that there is now scarcely a single competent general naturalist who is not prepared to accept some form of the doctrine of evolution." An illustrous example to the contrary is found in Professor Agassiz,—not to mention Chancellor Howard Crosby, or Professor Tayler Lewis (14).

If any one wishes to gratify at once his curiosity and his sense of beauty let him go up the great staircase of the British Museum, and, on entering the second corridor, turn either right or left to the cases which contain these sponges and other deep-sea forms—to which, by-the-by, in the present crowded state of the Museum, ghastly troops of monkeys serve as a foil—and there see how Nature is not only "*maxima in minimis*," greatest in her least, but "*pulcherrima in abditis*," fairest in her most hidden works; and how the Creative spirit has lavished as it were, unspeakable artistic skill on low-organized forms, never till now beheld by man, and buried not only in foul mud, but in the unsightly mass of their own living jelly.

But so it was from the beginning; and this planet, with its complicated wonders and beauties, was not made for man alone. Countless ages before man appeared on earth, the depths of the old chalk ocean teemed with forms as beautiful and as perfect as those, their lineal descendants, which the dredge now brings up from the Atlantic sea-floor (15).

This account, based mainly on contemporary quotations, perhaps provides some indication of the true beginnings of oceanography seventy-five years ago. It has been said that oceanography is a young and inclusive science (16), and the implication is that it has undergone rapid development. That the growth of this field is not universally appreciated is indicated by a review (not to be taken seriously) of Sverdrup, Johnson and Fleming (3), which states that: ". . . great advances have been made in the science, particularly with regard to the topography of the ocean beds, the chemistry of the sea and, what is probably the most astonishing of all, the investigations of life at great depths. Contrary to expectations, it has been ascertained that life is possible at all depths . . . " (17).

Following the Challenger Expedition efforts were mainly bent towards the discovery and description of new deep-sea forms; it was not until the 20th century that the science of oceanography began to take form in its modern and broad sense of the word. Under these circumstances this "Festschrift" is the more appropriate. The man for whom it is gathered together is pre-eminent in having advanced oceanography on two continents by his personal leadership and the world over by his researches and publications during the past three and a half decades.

REFERENCES

(1). Thomson, C. W. 1873. The Depths of the Sea. Macmillan and Co., London. XX + 527 pp.

(2). Herdman, W. A. 1932. Founders of Oceanography and Their Work. Edward Arnold and Co., London. XII + 340 pp.

(3). Sverdrup, H. U., M. W. Johnson and R. H. Fleming. 1942. The Oceans, Their Physics, Chemistry, and General Biology. Prentice-Hall, Inc., New York. X + 1087 pp.

(4). The Saturday Review, London, *35*, March 15, 1873: 352–354.

(5). Proc. Linn. Soc. Lond., Session 1881–1882: 67–68.

(6). Thomson, C. W. 1877. The Voyage of the Challenger. The Atlantic. Macmillan and Co., London. Vol. I: XXIX + 424; Vol. II: XIV + 396 pp.

(7). Moseley, H. N. 1879. Notes by a Naturalist on the "Challenger." Macmillan and Co., London. XVI + 620 pp.

(8). Dec. 5, 1872: 3.

(9). Dec. 28, 1872: IV.

(10). Chambers's Journal, London and Edinburgh, Vol. IV, No. 10 (2), 1873: 229–232.

(11). The British Quarterly Review, London, *57*, Jan. and Apr. 1873: 537–540.

(12). The Athenaeum, London, No. 2384, July 5, 1873: 18.

(13). The Guardian, London, June 25, 1873: 845.

(14). New-York Daily Tribune, April 1, 1873.

(15). The Times, London, April 11, 1873: 5.

(16). Bigelow, H. B. 1931. Oceanography, Its Scope, Problems, and Economic Importance. Houghton Mifflin Co., Boston. VI + 263 pp.

(17). Machinery Lloyd, June 24, 1944. (England)

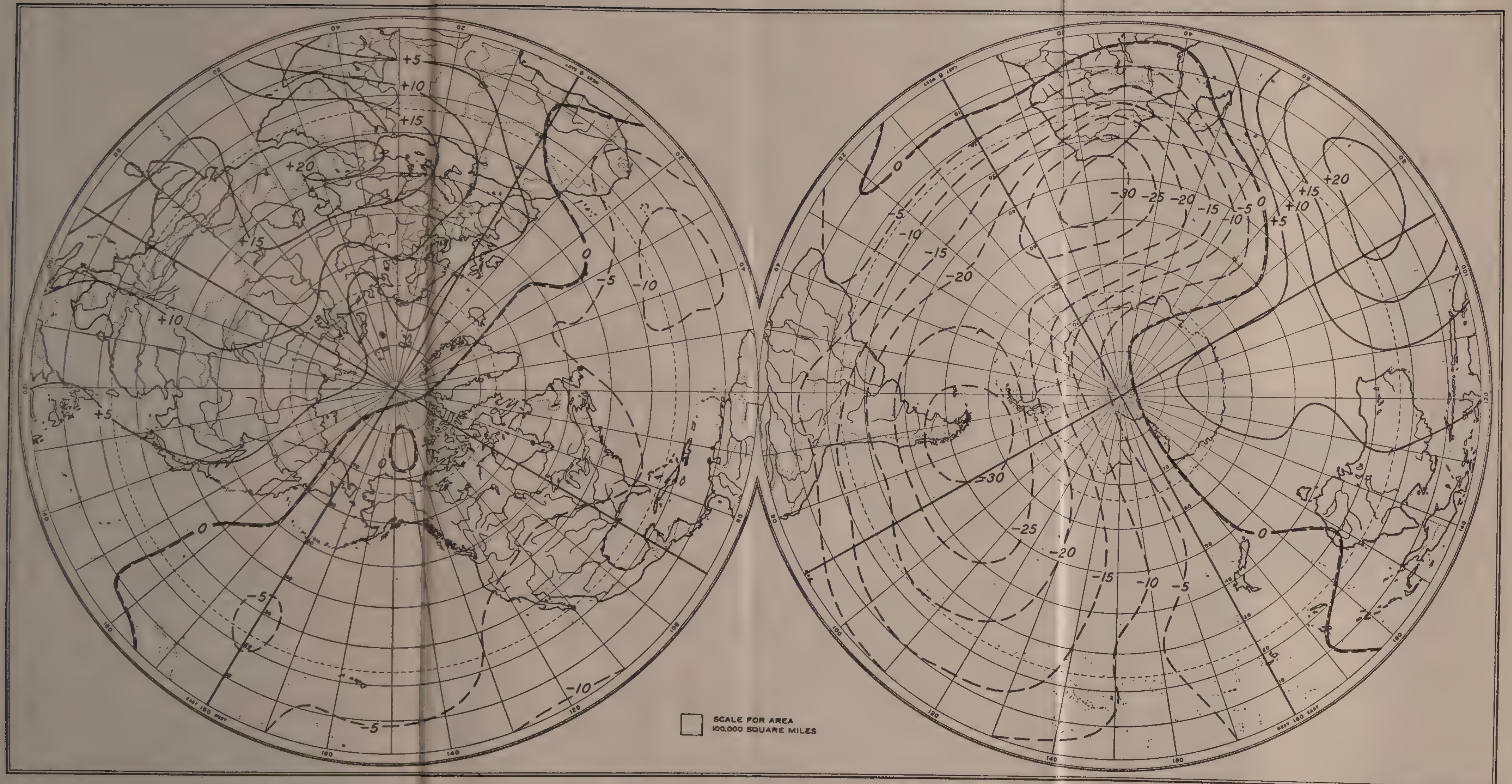

+5
+10
+15
+20
0
-5
-10
-15
-20
-25
-30
SCALE FOR AREA
100,000 SQUARE MILES

Figure 1. Lines of equal ratio of annual changes in total geomagnetic intensity to value of total intensities ($\Delta F/F$) for epoch 1942.5, expressed in units of 0.0001; full lines show positive values, dashed lines negative values.

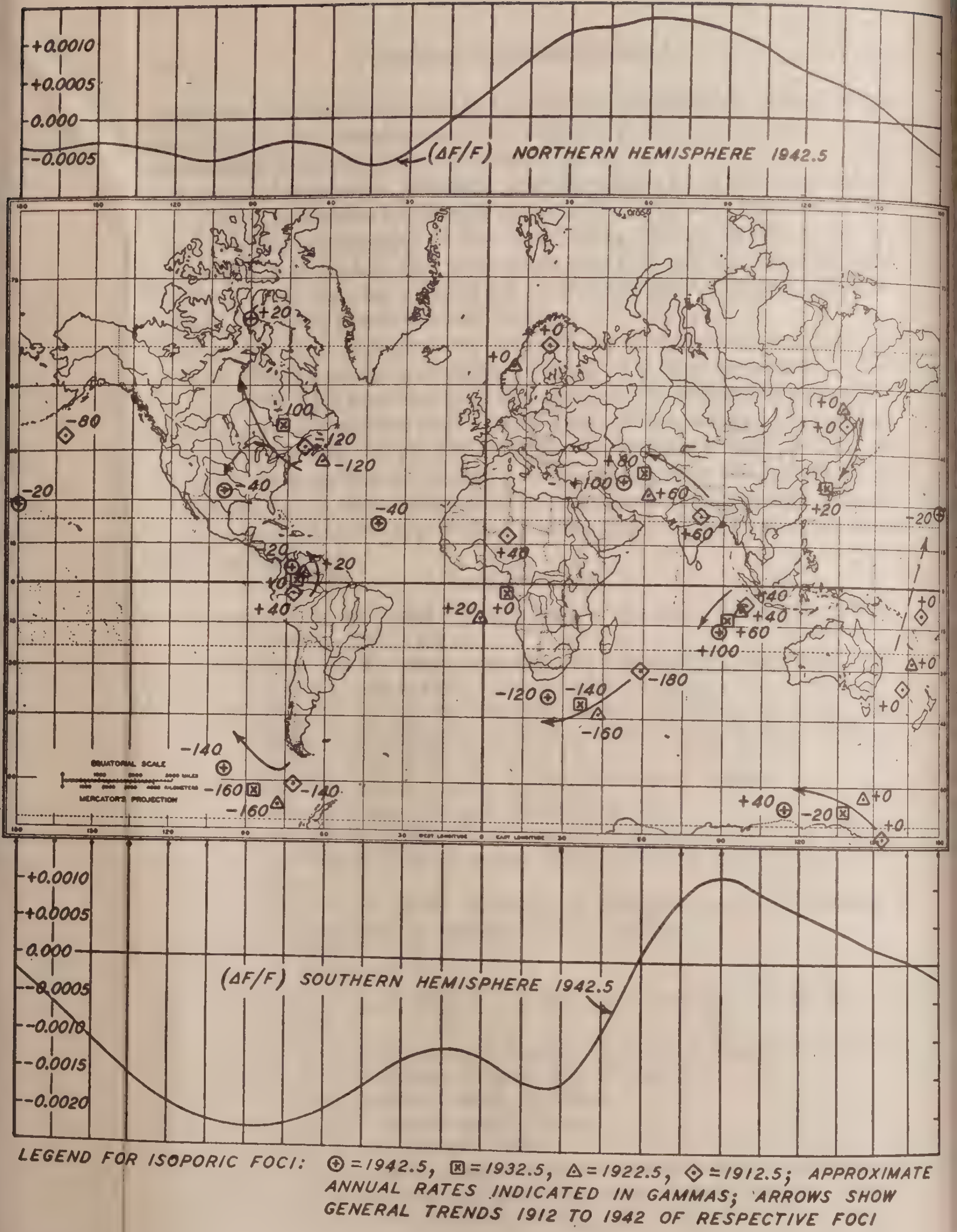

+0.0010
+0.0005
0.000
-0.0005
(ΔF/F) NORTHERN HEMISPHERE 1942.5
EQUATORIAL SCALE
MERCATOR'S PROJECTION
+0.0010
+0.0005
0.000
-0.0005
-0.0010
-0.0015
-0.0020
(ΔF/F) SOUTHERN HEMISPHERE 1942.5
LEGEND FOR ISOPORIC FOCI: ⊕=1942.5, ⊠=1932.5, △=1922.5, ◇=1912.5; APPROXIMATE ANNUAL RATES INDICATED IN GAMMAS; ARROWS SHOW GENERAL TRENDS 1912 TO 1942 OF RESPECTIVE FOCI

Figure 2. Isoporic foci of geomagnetic intensity for epochs 1912.5, 1922.5, 1932.5, and 1942.5 and general trends with time; curves of values of ($\Delta F/F$) for epoch 1942.5 for northern and southern hemispheres as averaged for 10° quadrilaterals of latitude and longitude from 80° north to 80° south.

OCEANOGRAPHY AND GEOMAGNETISM

BY

J. A. FLEMING

Carnegie Institution of Washington

Because of the large portion of the Earth that is covered by the oceans, the determination of accurate values of the geomagnetic elements at sea is a major objective in a world survey. It was not until 1905–1929 that realization of this objective was obtained through the systematic oceanic surveys sponsored by the Carnegie Institution of Washington through its Department of Terrestrial Magnetism. Earlier magnetic surveys at sea included those of the PARAMOUR (1698–1700), the EREBUS, the TERROR and the PAGODA (1840–45), the NOVARA (1857–60), the CHALLENGER (1872–76), and the GAZELLE (1874–76). Observations were also made by naval services of various countries and more lately by the DISCOVERY, the GAUSS, and by vessels of other Antarctic Expeditions.

All of the observations through 1900 were of varying degrees of accuracy determined by available instruments and by the disturbing factors inherent in the magnetic character of the vessels. The distribution of the observations, both as regards position and epoch, was inadequate to readily yield co-ordinated surface charts that applied to definite epochs, and the observations were frequently limited to but one or two elements. The CARNEGIE (1909–29) was a specially designed nonmagnetic vessel, equipped with newly designed instruments capable of yielding results of high accuracy under the difficult conditions inherent in work at sea. The CARNEGIE and her special equipment were developed following earlier experience (1905–08) on the brigantine GALILEE.

It has long been realized that irregularities in the surface distribution of the geomagnetic field over continental areas may bear definite relation to geologic features of the Earth's crust. In comparatively recent years geomagnetic data have been intensively utilized as an aid in the determination of the character of geologic substructure. A considerable part of the observed Earth's geomagnetic field may be ascribed to a uniformly magnetized sphere. The difference between this uniform or normal magnetization and the whole observed field, that is, the residual field, reflects those irregularities caused by geologic formations. The Earth's outer crust or shell is not homogeneous and is not uniform in its magnetic behavior. There are regions of local magnetic disturbance, many of which are caused by magnetic ore-deposits, some of them so great as to give rise to local poles and other irregularities of intensity twofold or even threefold the normal value in those regions. The anomalies at Kursk, Russia, at Berggiesshübel, Germany, and in New South Wales are striking examples of magnetic irregularities or anomalies. The great anomaly in South Africa resulting from the huge Pilansberg system of Paleozoic volcanic dykes shows strong magnetization in a direction opposite to the present-day geomagnetic field.

The magnetic anomalies over continental areas must have counterparts in the great oceanic basins. Therefore, the study of the former must aid investigation of oceanic structure. Indeed, observations on isolated islands and island groups almost invariably indicate abnormal magnetic conditions, and the magnetic survey of the oceans by the Carnegie Institution of Washington since 1905 has revealed similar abnormal conditions at sea.

The application of surface magnetic-survey results to the investigation of crustal features of the ocean bottom depends, however, both on differences in substructural physical properties and upon the magnitude of these differences. The magnetic susceptibility of rocks, some of which are strongly magnetic (as iron-ores and igneous and metamorphic rocks) and others weakly magnetic (as most sedimentary rocks) are important factors. Thus, observed surface magnetic anomalies afford a general reconnaissance method that will indicate regions that might be subjected to more intensive geophysical studies. The magnetic approach lacks depth-control; geologic features must increase in proportion to depth if they are to be revealed by geomagnetic data. Unless the magnetic anomaly observed at sea is over shallow water (or, if over deep water, unless it is of enormous proportions), no immediate exact information regarding particular areas of bottom substructure may be obtained by study of only those charts showing surface isomagnetic contours.

The geomagnetic field undergoes changes with time, that is, secular variation; the investigation of this phenomenon offers large potential value in the study of the Earth's crust. Herein is the great need for a continuance of the accurate surface magnetic survey of the oceans to determine the character of such changes with epoch and to determine the accelerations of such changes at various places.

From the data of several centuries, investigators have noted the apparent regularity of secular variation when considering only a single station or a small part of the Earth's surface, such as western Europe—say one-hundredth of the Earth's surface. When the observational data over the whole Earth are taken into account, however, it is found that secular variation is not a phenomenon that may be explained by periodic movements of the Earth's magnetic poles and the like; rather, it is a regional one connected with the large geological structure—oceans and continents—of the deeper layers of the Earth's crust. Techniques for the organization and interpretation of the data must be developed and experimental research must be conducted along lines which will supply information on basic physical problems related to this subject.

Surface secular-variation data are compiled from isomagnetic charts for various epochs. For succeeding epochs, shifts of the isomagnetics occur. Lines or contours joining points of equal annual change are called isopors. The foci of rapid isoporic changes raise questions of broad geophysical significance. Surface isopors for geomagnetic total intensity for the epoch 1942.5, based on the recently-published intensive analysis by the Carnegie Institution of Washington (data obtained during 1905–1945), are shown in Fig. 1. A large part of these changes may have origin in influences impressed by forces deep within the Earth. The rise and fall of the rate of secular change, and the slow expansion followed by the gradual contraction of the surface areas within which there have been excessive alterations in the magnetic elements, are significant of such changes. The distribution of isoporic foci is noteworthy. They are practically all in or near the great land-masses. Those foci found in the Pacific Ocean are generally of moderate intensity and not well defined, as shown in Fig. 2. These relations to the surface structure of the Earth indicate a causal relation with crustal or subcrustal movements and conditions.

Attention may be called to the apparent diminution of the intensity of the Earth's magnetic field which is marked over oceanic areas, especially in the western and southern hemispheres.

Observed earthquake-wave velocities and reflections for different regions and depths give some evidence that the crustal layer is thin under the Pacific Ocean. Under the Atlantic and Indian oceans the crustal layers are of appreciable depth. Under the Pacific Ocean different geological and geophysical properties from those found elsewhere may be expected. Thus, longitudinal distribution of surface isoporic foci agrees with that of land-areas as shown by the modest rates of annual change over the Pacific as compared with those over the Atlantic and adjoining continental areas. A definite control is necessary for a number of epochs to facilitate the investigation of causes producing and governing these surface progressive changes; accurate knowledge of the accelerations and of rates of annual distribution for different epochs is necessary.

Further data on the correlation between the surface-distribution of secular-change activity promise to give conclusions concerning both the character of the layer and the secular-variation processes localized in the crustal layer. Continued magnetic surveys at sea should amalgamate seismic and magnetic, and possibly gravimetric, approaches to crustal adjustments.

There is an interesting possibility that information on secular-variation may be extended to times more remote than those for which observations exist. When lava cools and freezes following a volcanic outburst it takes up a permanent magnetization dependent upon the orientation of the geomagnetic field at that given time. This, because of small capacity for magnetization in the Earth's field after freezing, may remain practically constant; thus the direction of the originally acquired permanent magnetization can be determined by laboratory tests, provided orientation of the mass to be tested is carefully noted and marked when it is removed.

Just as lavas, cooling through the Curie-point in the presence of a magnetic field, become magnetized in the direction of the field, so sediments containing microscopic magnetic particles also assume the direction of the prevailing magnetic field. This is "fossil" magnetization for the magnetic field of the epoch of deposition maintained in the sediment. Whether this fossil magnetism remains as a connate property of the rock, or is destroyed by subsequent events, depends upon the postnatal experience of the sediment. In so far as it remains, it affords a record of geomagnetism and offers an important clue to the history of the rock.

The project to measure geomagnetism at great depths in the sea may now be consummated by magnetic examination of bottom core-samples. Special equipment, already designed and tested, permits detection of magnetic moment equivalent to that of a good steel magnet one millimeter long and one-hundredth millimeter in diameter.

Tests on cores show distinct differences in the direction and intensity of magnetization for succeeding sections of cores. By interrelating individual cores through similarities in their magnetism, it may be possible ultimately to determine more accurately than by surface observations the extent and character of the regional magnetic anomalies over the oceans.

The excess of gravity in the oceans and a sudden decrease at the continental borders, with resulting deviations of isostasy, arise not only from sources in the Earth's crust but from far within the plastic layers below, thus creating current-systems below the crust—possibly downward below the oceans and upward below the continents. The surface chart of isoporic foci and their motions with time certainly lend strength to the idea which indicates that the interior of the Earth is more mobile than the external layers, not only as a whole but regionally. Interior currents may convey deep material with special magnetic properties, thus accounting for disturbances at the crust.

Therefore, it is important that improved knowledge of magnetic anomalies and of interrelations with gravimetric anomalies and seismic and tectonophysic conditions be obtained.

While more information on secular-variation changes in the Earth's magnetism is required for navigation, yet future geomagnetic data over the oceans are far more necessary to advance theoretical studies. It is of first importance to continue the work of the CARNEGIE, which was lost in 1929, because further surface surveys of like accuracy will enhance the theoretical value of the work already done. Investigation on diminution of the geomagnetic field—so pronounced in the southern and western hemispheres, as already stated—calls for more data over the oceans.

Theoretical investigations demanding continuation of oceanic geomagnetic surveys include, among others, the following:

(*a*) Determination of secular-variation of progressive changes of the geomagnetic field involving their accelerations, which the data accumulated so far indicate cannot be extrapolated reliably over periods as long as five years. A definite control is necessary for a number of epochs to facilitate the investigation of causes producing and governing these progressive changes.

(*b*) The study of regions of local disturbance, particularly of those indicated by earlier work over "deep-sea" areas, including accompanying determination of oceanic depths, of gravity and of the study of core-samples.

(*c*) The determination of additional distribution-data in a few large areas not already covered.

The question arises whether the theoretical requirements might not be met in a less expensive way than through construction and maintenance of nonmagnetic vessels. Study has indicated what might be done to control magnetic secular-variation data over the oceans between Lat. 60° N. and 60° S. through observations on land only. The maximum control which could be so effected on the surface would result from some 150 secular-variation stations along the coasts of the continents and on islands. Some 90 of the stations are readily accessible. The remainder include the more inaccessible islands which are subject, generally, to magnetic local disturbance. These islands introduce uncertainties both in the effects upon secular-variation changes and in the relation between the normal and the island value, even though the inaccessibility of stations insures exact reoccupations. The reduction to common epoch would be more difficult because of the length of intervals between reoccupations and of the lack of the better distribution of data to be had from observations at sea.

Even if the complete scheme for control by observations on land could be attained, it appears that the areas lacking necessary data would be very large. These areas approximate nearly 15,000,000 square miles in the Pacific Ocean, over 3,000,000 in the Atlantic Ocean, and some 6,000,000 in the Indian Ocean—a grand total of 24,000,000 square miles in the oceanic areas between parallels 60° N. and 60° S. Local disturbances at many stations on islands, which would undoubtedly result in some data being unsuitable for discussion, actually would make for an even greater total area. These areas involve portions of the Earth's surface where there are at present the greatest irregularities in the progressive character of the secular variation in the Central and South Atlantic, Indian, North Pacific, East Central Pacific, and South Pacific oceans. Data of like kind over the remaining one-seventh of the Earth in the Polar areas may be expected through special expeditions on land and in the air.

Because of the great need for continuing the operations conducted for a quarter-century by the Carnegie Institution of Washington, and in view of this Institution's decision not to replace the CARNEGIE, the British Admiralty designed and completed the nonmagnetic vessel RESEARCH. Already there are marked uncertainties in present isomagnetic charts because of rapid isoporic changes in certain regions. Therefore, cruises by the RESEARCH are urgent.

The task of the geophysical surface survey of the oceans is so great that the hydrographic services of other nations should provide similar vessels with equipment and personnel to take an appropriate share in the execution and in the co-ordination of such service.

The present slow and costly methods of surface geomagnetic surveys, both on land and sea, may now be supplemented by new techniques and methods. The procedure of measuring the Earth's field at different times in different places in one region during one or more years, and in another at a different epoch, is inadequate for all the needs. What is required is a description of the geomagnetic field during a given year, based on measurements made during that year at an orderly set of points sufficiently close together and so spaced as to be nearly independent of topography of areas of land and sea, and of climate. That such measurements can be made from planes is now feasible instrumentally. No matter how accurately the magnetic elements are determined at a given surface point, information is still inadequate for surface isomagnetic maps in the absence of observations at intervening points. For a successful program an almost infinite number of stations is required.

This condition may now be remedied by magnetic observations with instruments airborne above the surface of the Earth, thus giving observations that are farther removed from magnetic anomalies which depreciate the accuracy of surface observations. By means of instruments carried at high speeds, continuous observations may be obtained over a large area in a single day, so that in a relatively short time a survey of the normal geomagnetic field can be made. The accuracy of such techniques can be kept at a standard that is more than adequate for both practical and theoretical purposes.

Thus, as a result of an effective airborne instrument (*a*) large areas may be covered quickly and continuously, (*b*) surface anomalies, which are of no significance in the preparation of charts of the normal field, may be eliminated (by flying at altitudes of, say 20,000 feet, smoothed values of the geomagnetic elements are obtained directly, without the necessity of mathematical treatment or judgment in processing the data) and (*c*) flights in two or more horizontal planes may define uniquely the magnetic field.

Perfection of a suitable airborne instrument will require considerable time and funds; nevertheless, the end results would justify large expenditures of both, particularly for observations over the oceans where the elimination of surface anomalies at high levels will make for better interpretation of surface irregularities over the normal field.

In conclusion, the facts of the geomagnetic research in the field and in the laboratory must contribute much to formulate a complete picture of the Earth's crust and particularly of that great portion which is covered by the oceans.

ATMOSPHERIC TIDES[1]

BY

JACOB BJERKNES

University of California, Los Angeles

The subject of atmospheric tides has been and still is a great challenge to the theoretical meteorologist. Despite the many brilliant contributions to atmospheric tidal theory from the time of Laplace up to the present, one of the essential tidal features—the surprisingly great amplitudes of the semidiurnal tide produced by the sun—still calls for a really convincing explanation. This article does not bring any refinements in existing tidal theories, but it does provide the suggestion that the model to which the theories have hitherto been applied is not sufficiently realistic. With a more true-to-nature model of the semidiurnal atmospheric tide, there seems to be a possibility for a simple explanation of the great tidal amplitude.

Fig. 1 summarizes the empirical facts about the solar harmonic components W_2^2, Z_2 and W_4^3 as observed in the sea-level atmospheric pressure. Note the dominating character of the sectorial progressive 12h wave, W_2^2, with its maximum pressure oscillation of ± 1.3 mb at the equator. The standing 12h oscillation, Z_2 and the tesseral progressive 8h oscillation W_4^3 are, in comparison, quite weak. A global 24h oscillation also exists, but it is so mixed up with the thermal sea- and land-breeze and with valley-mountain circulations of 24h period that it proves difficult to isolate. The evidence from pinpoint islands in the tropical Pacific without local 24h wind systems is that the global diurnal tide has an amplitude of the order one-fifth of that of the semidiurnal tide.

All lunar tidal components in the atmosphere are much weaker than the corresponding solar ones, and it is only the sectorial progressive wave, W_2^2, that has been found empirically (Chapman, 1918). The observed lunar tide pressure oscillation amounts to ± 0.09 mb at the equator, with a decrease towards higher latitudes in the same fashion as shown for the solar W_2^2 in Fig. 1.

The contrast to the ocean tides is striking. In the oceans the moon produces a stronger tidal component than the sun because of the greater tidal force exerted by the moon. In the atmosphere that relation is reversed, although the ratio between lunar and solar tidal forces must be the same. The classical explanation of this phenomenon is based on the assumption of resonance. The atmosphere is

---

[1] U. C. L. A. Department of Meteorology, Papers in Meteorology, No. 11.

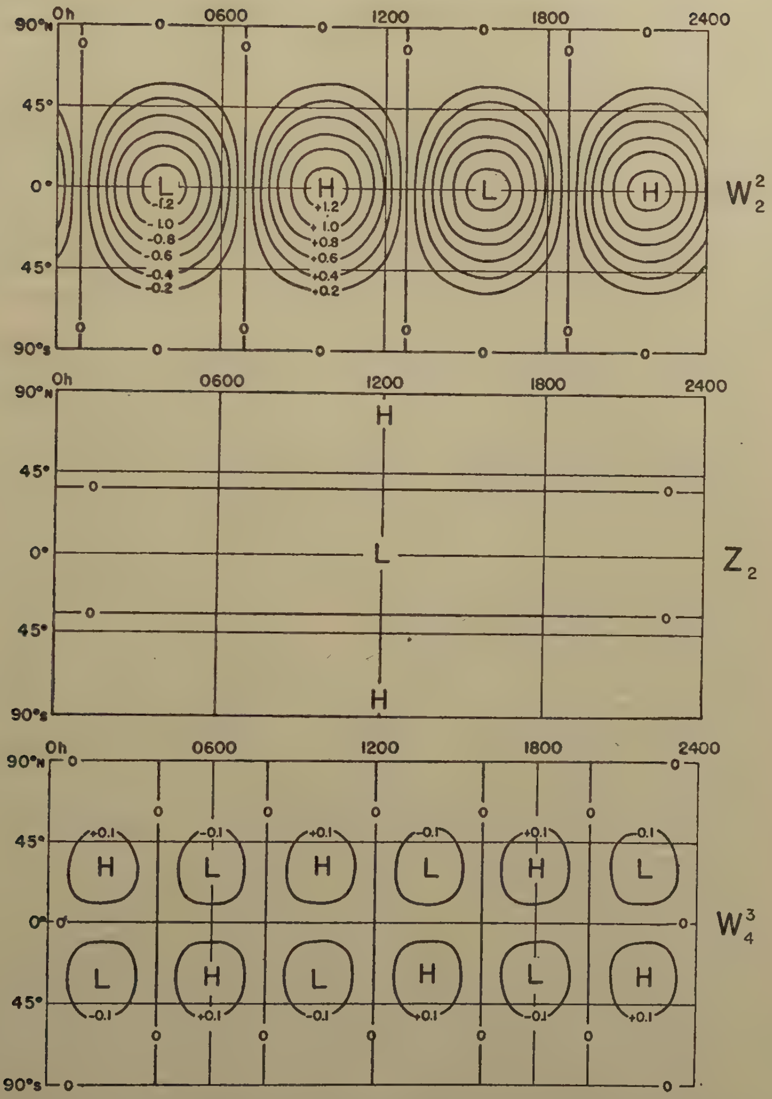


Figure 1. The planetary harmonic components of atmospheric tides in solar time as observed at sea level: W_2^2, the progressive sectorial 12h oscillation (pressure interval 0.2 mb); Z_2, the standing 12h oscillation, depicted at 11h or 23h Greenwich time (pressure amplitude about 0.06 mb at the equator and 0.12 mb at the poles); W_4^3, antisymmetric part of the progressive 8h oscillation averaged from November to February (pressure interval 0.1 mb). From Bartels, 1939.

supposed to have a period of proper oscillation very close to 12 solar hours (in fact, within less than 4 minutes), so that the semidiurnal tide can be amplified by resonance. It has been impossible to prove by dynamic reasoning that the atmosphere actually is tuned in its proper oscillations as exactly as the classical resonance theory requires. The temperature and pressure parameters of the oscillating atmosphere vary greatly with season, as well as from the one hemisphere to the other; nevertheless, the amplitude of the semidiurnal tide at each locality varies by a few per cent only (maximum at both equinoxes). It would be a great relief for the resonance theory if it could be shown that the tuning of the proper oscillations of the atmosphere is not required to be as sharp as indicated above.

It has long been known that the semidiurnal tide has some lag in phase with height, but, to my knowledge, the implications of that feature for the dynamics of the tide have not been expressed in tidal theory. Fig. 2 shows a profile of the "tilting 12h tide" as obtained by Wagner (1932, 1938) through harmonic analysis of pressure records at elevations from 522 m to 4,575 m in the Alps. At the highest elevation the phase of the semidiurnal tide is about two hours behind that observed globally at low level stations. The delay of the pressure extremes with height is equivalent to an eastward tilt of the westward moving crests and troughs.

The temperature field of the tilting tide was determined also by Wagner on the basis of the pressure oscillation and with the assumption of quasistatic equilibrium (neglect of vertical acceleration and vertical Coriolis force). The temperature profile is shown in the lower part of Fig. 2. The semidiurnal temperature wave (obtained indirectly from the much more accurate determination of the pressure wave) is also delayed with height, but the time lag does not increase beyond the elevation of 2 km above sea level.

The physical process responsible for the 12h temperature wave can be read from Fig. 3, which contains the 24h and 12h terms of the harmonic analysis of temperature at the various heights covered by Wagner's Alpine data. The temperature curves again refer to the temperature derived indirectly through the analysis of the pressure data at mountain locations. In all levels the 24h temperature wave has a greater amplitude than that of the 12h wave and represents in a first harmonic approximation the day and night differences of temperature. The 12h wave of temperature is of such a phase that it changes the 24h sine function into one with a steep temperature increase in the morning hours and a more gentle decrease of temperature in the afternoon. The physical process involved, both in the 24h and in the 12h temperature wave, is obviously that of radiation, which calls for a

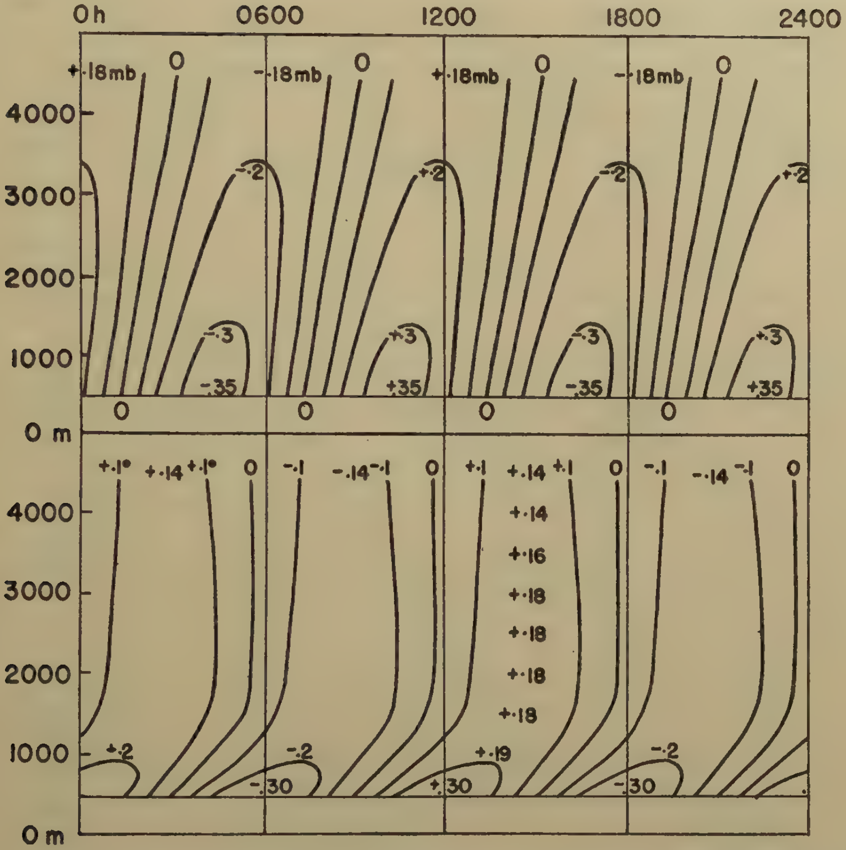


Figure 2. *Upper part:* Profile of the semidiurnal harmonic component of pressure, based on data from Alpine summit stations. *Lower part:* Profile of the semidiurnal component of temperature derived from corresponding pressure data with the assumption of quasistatic conditions. (Data from Wagner, 1932.)

higher temperature in the daytime than at night; more specifically, radiation also calls for a rapid rise in temperature after sunrise and a slower decrease of temperature after the diurnal maximum. The "sunrise effect," known from the shape of thermograms near the ground, is revealed also by the temperature curves at higher levels. In fact, the ratio of the amplitudes of the 12h and 24h temperature waves increases with height, indicating that the sunrise effect on the shape of the diurnal curve of free-air temperature becomes more accentuated with height.

With the assurance of the physical reality of the 12h temperature wave, we may now venture to extrapolate the semidiurnal waves of pressure and temperature upwards beyond the levels of the highest points of observation in the Alps. In the lower portion of Fig. 4 the crests and troughs of the semidiurnal pressure and temperature waves have been copied as solid lines from Fig. 2, and up beyond 4.5 km the dashed lines show our tentative extrapolation. The temperature

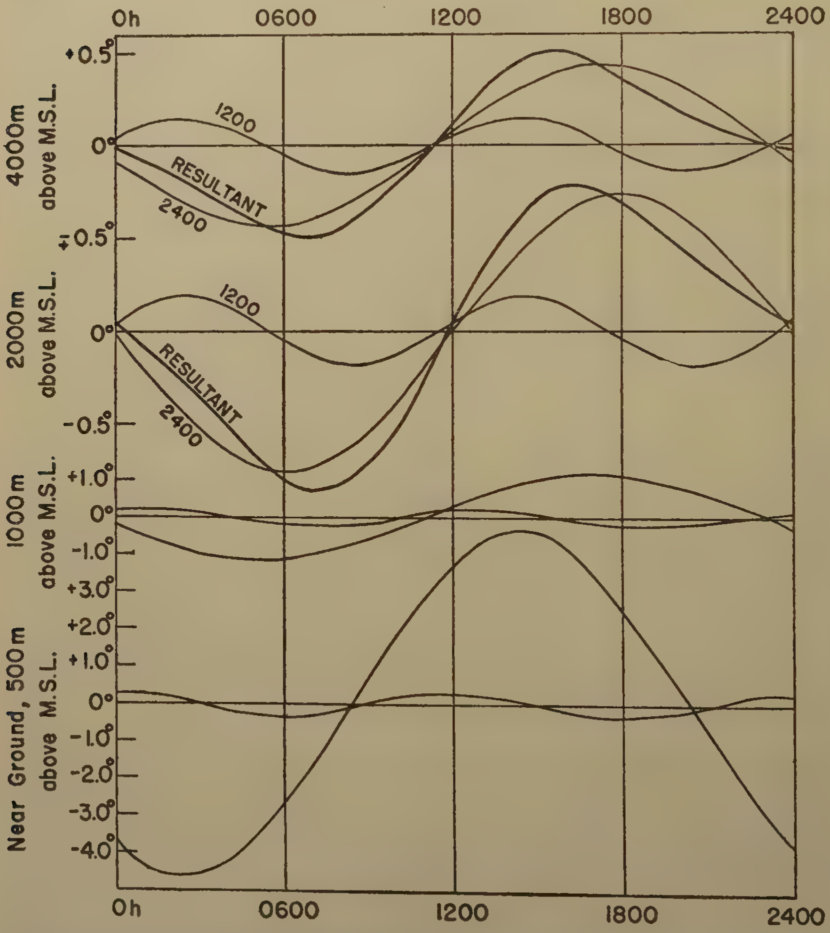


Figure 3. Diurnal and semidiurnal harmonic components of temperature, derived from harmonic analysis of pressure records at Alpine summit stations. (Data from Wagner, 1932.)

wave, which is timed by the "sunrise effect," continues quasivertically upward with an insignificant tilt due to the earlier arrival of sunrise with height. The pressure wave tilts asymptotically towards phase coincidence with the temperature wave, as prescribed by hydrostatics.

The total picture of the semidiurnal atmospheric tide is thus composed of pressure waves moving at the speed of local time from east

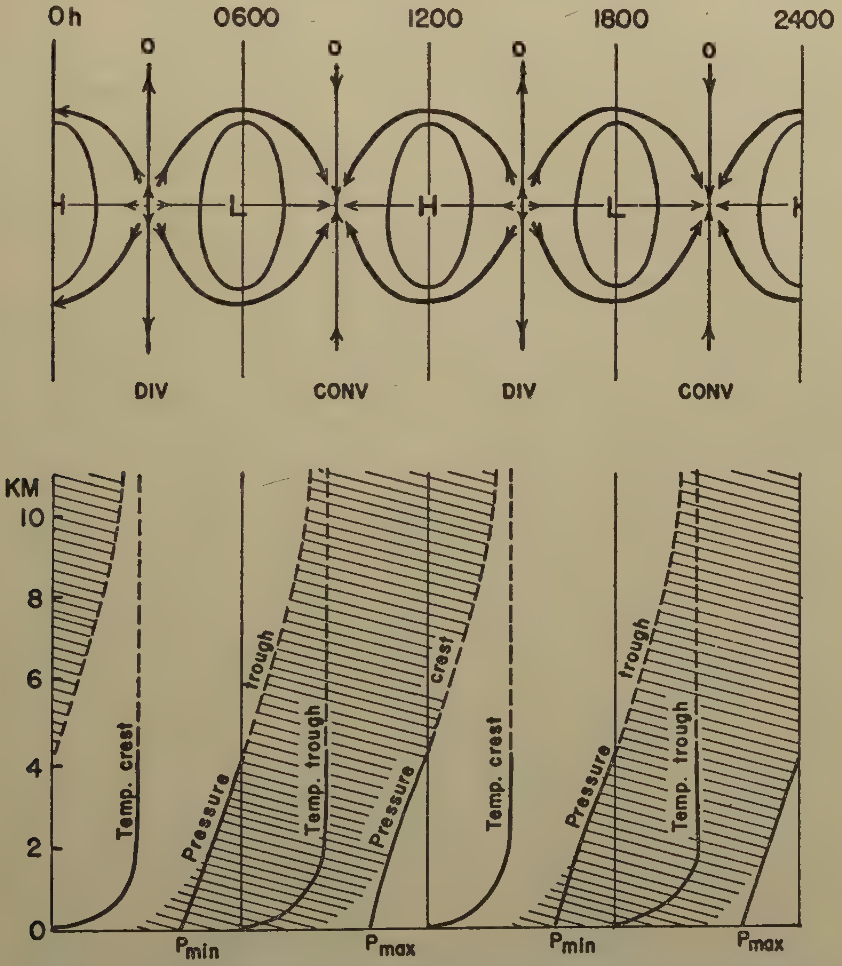

0h
0600
1200
1800
2400
0
0
0
0
H
L
H
L
H
DIV
CONV
DIV
CONV
KM
10
8
6
4
2
0
Temp. crest
Pressure
trough
Temp. trough
Pressure
crest
Temp. crest
Pressure
trough
Temp. trough
Pmin
Pmax
Pmin
Pmax

Figure 4. *Upper part:* Oscillation pattern of the semidiurnal tide with phase times as observed at roughly 4.5 km. *Lower part:* Profile of the tilting semidiurnal tide. Shaded space indicates convergence of the horizontal velocity field; unshaded space, divergence.

to west and tilting backwards with height all the way up to the stratosphere. At 4.5 km, about halfway through the atmosphere by weight, the phase times of the pressure waves are those of a simple semidiurnal gravitational tide produced by the sun. Below 4.5 km the pressure waves have earlier phase, and above 4.5 km later phase, than that of a simple gravitational tide because of the tilt produced by thermal asymmetry.

According to the theory of Pekeris (1937), the semidiurnal solar tide is of cellular structure and changes phase at about 30 km. That feature lies beyond the upper limit of Fig. 4 and will not influence the following discussion because of the small values of air density up beyond 30 km.

In the upper part of Fig. 4 a first approximation map of the horizontal projection of tidal streamlines has been inserted. It is visualized as applying to the level of roughly 5 km, where the observed tide has its pressure maxima at noon and at midnight. At other elevations the same oscillation pattern would be displaced as shown by the tilting crest and trough lines in the zonal profile. The pattern of tidal winds is assumed to be that derived by Bartels (1928); it is "antigeostrophic," that is, the pressure gradient and the Coriolis force act in the same direction and force the particles around in anticyclonic orbits with 12h period, and hence shorter than the inertia period 12h/sin φ.

The spaces of horizontal velocity divergence and convergence, limited roughly by the longitudes of maximum and minimum pressure, must tilt just as do the crests and troughs of pressure. That is indicated in Fig. 4 by shading in the zonal profile. In the surface layers, friction must modify the distribution of divergence to a certain extent. That effect can best be visualized at the equator, where the tidal wind is either east or west and where the force of friction can be assumed to be right opposite to the instantaneous wind. The tidal wind in the friction layer will reach its maximum when the force of friction and that of the pressure gradient are equal and opposite, which must occur some time during the phase when wind and pressure gradient have equal direction, which in turn means somewhere in the western half of the high and in the western half of the low pressure phase. Those places of maximum tidal wind, westward and eastward, will presumably indicate also the location of zero horizontal divergence of the tidal wind in the friction layer. In the zonal profile the zero lines of horizontal divergence of the tidal wind have been drawn accordingly.

The tendency equation applied to the semidiurnal pressure change at the ground (p = pressure, t = time, g = acceleration of gravity, ρ = air density, v_x and v_y = horizontal wind components),

$$(1)\quad \left(\frac{\partial p}{\partial t}\right)_o = -\int_o^\infty g\rho\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)dz - \int_o^\infty g\left(v_x\frac{\partial \rho}{\partial x}+v_y\frac{\partial \rho}{\partial y}\right)dz\,,$$

reduces to

$$(1')\quad \left(\frac{\partial p}{\partial t}\right)_o = -\int_o^\infty g\rho\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)dz\,,$$

the density advection term being only about 10^{-3} times the magnitude of the divergence term. We may therefore discuss the semidiurnal pressure change directly from Fig. 4, which contains the zonal profile of horizontal divergence, $\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)$. The pressure rise at the ground, centered around 7h and 19h local time, comes from the surplus of horizontal convergence in the vertical column, and the pressure fall, centered around 1h and 13h local time, comes from the surplus of horizontal divergence. The mechanism of the westward progression of the pressure oscillation lies in the alternation of convergence and divergence in the proper phase to give pressure rise west of the high and pressure fall west of the low. In order to give zero value for the pressure tendency at the ground at 4h and 16h local time, the upper divergence must just compensate the lower frictional convergence in the right hand integral in (1′), and analogously, at 10h and 22h the upper convergence must compensate for the lower divergence. Without the frictional divergence and convergence the pressure tendency would have been positive at the crest of the high and negative at the trough of the low, which would have meant a progressively increasing amplitude of the semidiurnal pressure oscillation. The great amplitude of the semidiurnal pressure oscillation can therefore be understood as an effect of its tilt in space, which in turn is the result of its thermal asymmetry. The semidiurnal pressure oscillation will maintain an adjusted intensity at which the frictional effect near the ground is just strong enough to check the growth of the wave motion.

An exact theory, built on the rough outline above, will probably show that a less rigorous approach of the period of proper oscillations to 12 solar hours is required for tilting pressure waves than for non-tilting ones. The big semidiurnal pressure amplitudes would then cease to be a theoretical enigma.

The 24h solar tide, which was also analyzed by Wagner, does not show any appreciable tilt of pressure crests and troughs and will therefore not be subject to the intensification described for the 12h solar tide. As for the lunar tide there is no reason to expect any thermal asymmetry with inherent tilt and intensification, and hence it is understandable that the lunar tide remains much weaker than the solar tide.

The tilt of the semidiurnal atmospheric tide has of course no analogy in the ocean, because the ocean has no semidiurnal temperature oscillation in depth. The theory of ocean tides, which in a general way has set the pattern for the theory of atmospheric tides, will call for a generalization to fit the tilting semidiurnal atmospheric tide.

REFERENCES

BARTELS, JULIUS

1928. Gezeitenschwingungen der Atmosphäre. Wien-Harms Handbuch der Experimentalphysik. Geophysik. Teil 1: 161–210, diagr. p. 206. Akademische Verlagsgesellschaft, Leipzig. 699 pp.

1939. Sonnen- und mondentägige Luftdruckschwankungen. Hann-Süring Lehrbuch der Meteorologie, pp. 276–306, diagr. p. 302. Willibald Keller, Leipzig. 480 pp.

CHAPMAN, S.

1918. The lunar atmospheric tide at Greenwich, 1854–1917. Quart. J. R. met. Soc., *44:* 271–280.

PEKERIS, C. L.

1937. Atmospheric oscillations. Proc. roy. Soc. London, (A) *158:* 650–671.

WAGNER, A.

1932. Der tägliche Luftdruck- und Temperaturgang in der freien Atmosphäre und in Gebirgstälern. (Gerland's) Beitr. Geophys., *37:* 315–344.

1938. Theorie und Beobachtung der periodischen Gebirgswinde. (Gerland's) Beitr. Geophys., *52:* 408–449.

ON DYNAMIC STABILITY OF ZONAL CURRENTS[1]

By

JÖRGEN HOLMBOE

University of California, Los Angeles

The problem of the dynamic stability of atmospheric currents has been investigated by many authors during the past 20 years. The present paper deals with certain aspects of this problem which was first treated by Solberg [1], Høiland [2], and Kleinschmidt [3], and more recently by Van Mieghem [4], and Fjørtoft [5].

The system under consideration is a steady zonal current. The zonal wind varies arbitrarily with latitude and height and is balanced by the proper baroclinic field of solenoids. The effects of friction are neglected. This system is known as a steady circular vortex and satisfies the following equation of (absolute) motion:

$$- \omega^2 \boldsymbol{R} = \boldsymbol{g} - \alpha \nabla p \,, \tag{1}$$

where ω is the absolute angular speed. The equation states that the centripetal acceleration in the steady circular motion, $- \omega^2 \boldsymbol{R}$, is the resultant of the force of gravitation, $\boldsymbol{g}$, and the pressure force, $- \alpha \nabla p$ The resultant of the force of gravitation and the centrifugal force will be called the apparent gravity and will be denoted by $\boldsymbol{g}^*$;

$$\boldsymbol{g}^* = \boldsymbol{g} + \omega^2 \boldsymbol{R} = \alpha \nabla p \,. \tag{2}$$

The steady circular vortex is then subjected to a small vortex ring perturbation, such that the meridional displacements are the same in all meridional planes. This perturbation will not destroy the axial symmetry of the field of solenoids, so the circulation of the expanding and contracting circles will remain unchanged. The physical changes of the fluid particles are assumed adiabatic, so the potential temperature will be individually conserved. Thus, if c denotes the velocity circulation per radian, and θ denotes the potential temperature, we have the two conservation laws,

$$\frac{dc}{dt} = 0 \,; \quad \frac{d\theta}{dt} = 0 \,. \tag{3}$$

The perturbed motion must satisfy the following equation:

$$\frac{\partial \boldsymbol{v}}{\partial t} - (\omega^2 + \overline{\omega^2}) \boldsymbol{R} = \boldsymbol{g} - (\alpha + \bar{\alpha}) \nabla (p + \bar{p}) \,, \tag{4}$$

---

[1] U. C. L. A. Department of Meteorology, Papers in Meteorology, No. 10.

where the bar over each symbol denotes the *local* variation of the corresponding quantity. (The radius vector $\mathbf{R}$ and the force of gravitation $\mathbf{g}$ have no local variation.) When the last term on the right in (4) is developed, and the second order term $\bar{\alpha}\nabla\bar{p}$ is neglected, we find from (1) and (4) the perturbation equation:

$$\frac{\partial \mathbf{v}}{\partial t} - \overline{\omega^2}\mathbf{R} = -\bar{\alpha}\nabla p - \alpha\nabla\bar{p}\,.$$

Introducing from (2) the apparent gravity in the first term on the right, the perturbation equation becomes

$$\frac{\partial \mathbf{v}}{\partial t} = \overline{\omega^2}R\nabla R - \frac{\bar{\alpha}}{\alpha}\mathbf{g}^* - \alpha\nabla\bar{p}\,. \tag{5}$$

The physical meaning of this equation becomes clear when the local variations of ω^2 and α are expressed in terms of the local variations of the conservative quantities c and θ. If $\mathbf{r}$ denotes the meridional displacement, the local variations of these latter quantities are from (3) given by

$$\bar{c} = -\mathbf{r}\cdot\nabla c\,; \quad \bar{\theta} = -\mathbf{r}\cdot\nabla\theta\,. \tag{6}$$

The local variation of ω^2 is obtained by performing local differentiation of the expression

$$\omega^2R^4 = c^2\,,$$

which gives

$$\overline{\omega^2}R^4 = 2c\bar{c} = 2\omega R^2\bar{c}\,,$$

or, when the expression for $\bar{c}$ is substituted from (6):

$$\overline{\omega^2}\mathbf{R} = -2\omega\frac{\nabla c}{R}\cdot\mathbf{r} = -2\omega q\left(\frac{\nabla c}{|\nabla c|}\cdot\mathbf{r}\right), \tag{7}$$

where q is the magnitude of the vorticity.[2]

The local variation of α is obtained by performing differentiation of the expression

$$\alpha = \frac{R\theta}{100}\left(\frac{p}{100}\right)^{\kappa-1}; \quad \kappa = c_p/R\,,$$

which gives

$$\frac{\bar{\alpha}}{\alpha} = \frac{\bar{\theta}}{\theta} + (\kappa - 1)\frac{\bar{p}}{p}\,,$$

---

[2] If ψ denotes the zonal angle we have:

$$\mathbf{q} = \nabla\times\mathbf{v} = \nabla\times vR\nabla\psi = \nabla c\times\nabla\psi\,;$$

therefore, $q = |\nabla c||\nabla\psi| = |\nabla c|/R\,.$

or, when the expression for $\bar{\theta}$ is substituted from (6):

$$\frac{\bar{\alpha}}{\alpha} = -\frac{\nabla\theta}{\theta}\cdot\boldsymbol{r} + (\kappa - 1)\frac{\bar{p}}{p}. \tag{8}$$

When (7) and (8) are introduced in the perturbation equation (5), it becomes

$$\frac{\partial \boldsymbol{v}}{\partial t} = -2\omega q\nabla R\left(\frac{\nabla c}{|\nabla c|}\cdot\boldsymbol{r}\right) + \frac{|\nabla\theta|}{\theta}\boldsymbol{g}^*\left(\frac{\nabla\theta}{|\nabla\theta|}\cdot\boldsymbol{r}\right) - \alpha\left\{\nabla\bar{p} + (\kappa - 1)\,\bar{p}\,\nabla\ln p\right\}. \tag{9}$$

To facilitate the further discussion of this equation we introduce the following abbreviations:

$$\boldsymbol{a} = 2\omega q\nabla R, \tag{10}$$

$$\boldsymbol{a}' = \nabla c/|\nabla c|, \tag{11}$$

$$\boldsymbol{b} = -\frac{|\nabla\theta|}{\theta}\boldsymbol{g}^* = -\frac{g^*}{\theta}\frac{\partial\theta}{\partial n_\theta}\frac{\nabla p}{|\nabla p|}, \tag{12}$$

$$\boldsymbol{b}' = \nabla\theta/|\nabla\theta|, \tag{13}$$

$$\boldsymbol{F}(\bar{p}) = \alpha\left\{\nabla\bar{p} + (\kappa - 1)\,\bar{p}\,\nabla\ln p\right\}. \tag{14}$$

With these notations introduced in the perturbation equation (9), it takes the final form

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\boldsymbol{a}\,(\boldsymbol{a}'\cdot\boldsymbol{r}) - \boldsymbol{b}\,(\boldsymbol{b}'\cdot\boldsymbol{r}) - \boldsymbol{F}(\bar{p}). \tag{15}$$

The equation gives the meridional acceleration in the disturbed motion as a resultant of three forces, each of which has a simple physical interpretation; see Fig. 1.

The first force on the right in (15) represents the centrifugal buoyancy. It is directed opposite to the vector $\boldsymbol{a}$ (10) which has the direction of the centrifugal force and the magnitude $2\omega q$. This magnitude measures the inertial stability of the vortex. It is equal to the square of the frequency of pure inertial oscillations [6]. The vector $\boldsymbol{a}'$ (11) is a unit vector perpendicular to the vorticity surfaces. Accordingly, only displacements perpendicular to these surfaces are active in the generation of centrifugal buoyancy. If the vortex is irrotational ($q = 0$), the vector $\boldsymbol{a}$ is zero, and the vector $\boldsymbol{a}'$ becomes indeterminate. All centrifugal buoyancy vanishes in the irrotational vortex.

The second force on the right in (15) represents the gravitational buoyancy. It is directed opposite to the vector ***b*** (12), which has the direction of the pressure gradient and the magnitude $\frac{g^*}{\theta}\frac{\partial \theta}{\partial n_\theta}$. This magnitude measures the gravitational stability of the vortex. It is equal to the square of the frequency of pure gravitational oscillations. The vector ***b'*** (13) is a unit vector perpendicular to the isentropic surfaces. Accordingly, only displacements perpendicular to these surfaces are active in the generation of gravitational buoyancy. If the vortex is isentropic (θ = const), the vector ***b*** is zero and the vector ***b'*** becomes indeterminate. All gravitational buoyancy vanishes in the isentropic vortex.

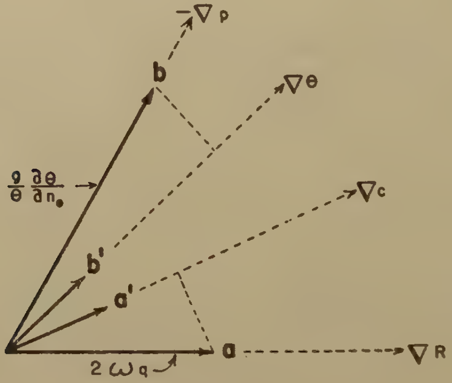


Figure 1.

The third force on the right in (15) represents the buoyancy caused by the local pressure variations. This term will not be discussed further in this paper.

The stability of the circular vortex is determined by the direction of the meridional acceleration (the resultant buoyancy) with reference to the meridional displacement. The stability criterion may be formulated as follows:

If for *all* displacements the corresponding resultant buoyancy has a component opposite to the displacement, the vortex is stable. If, on the other hand, there exist displacements for which the resultant buoyancy has a component in the direction of the displacement, the vortex is unstable. The problem of stability is therefore equivalent to that of finding the geometric distribution of the resultant buoyancy as a function of the meridional displacement.

When the problem is confined to the determination of the stability criterion only, it is customary to neglect the buoyancy caused by the local pressure variations. This is equivalent to using the so-called parcel method in the investigation of hydrostatic (gravitational) stability of a resting atmosphere. The physical justification for this simplification is probably that the local pressure variation merely weakens the action of the dynamic buoyancies without altering their direction or sign.

We shall accordingly, for the purpose of investigating the stability of the vortex, neglect the last term on the right in the perturbation equation (15). The remaining two terms on the right can then be written in the following compact form:

$$\frac{\partial \boldsymbol{v}}{\partial t} = -(\boldsymbol{a}\boldsymbol{a}' + \boldsymbol{b}\boldsymbol{b}') \cdot \boldsymbol{r}. \tag{16}$$

Mathematically, the resultant buoyant force appears here as a linear two-dimensional vector function of the displacement $\boldsymbol{r}$. The four vectors $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{b}$, $\boldsymbol{b}'$, which define the vector function, are not independent: In the undisturbed vortex the baroclinic field of solenoids is balanced by the proper shear in the windfield. The mathematical expression for this balance of the mass field by the so-called thermal wind is obtained by applying curl-differentiation to equation (1). Doing this, and using substitutions similar to those in equations (7) and (8), we find, after a simple calculation,

$$\boldsymbol{a} \times \boldsymbol{a}' + \boldsymbol{b} \times \boldsymbol{b}' = 0. \tag{17}$$

Thus the rotation $\boldsymbol{a} \rightarrow \boldsymbol{a}'$ is opposite to the rotation $\boldsymbol{b} \rightarrow \boldsymbol{b}'$, and the component of $\boldsymbol{a}$ perpendicular to $\boldsymbol{a}'$ is equal to the component of $\boldsymbol{b}$ perpendicular to $\boldsymbol{b}'$. The vectors in Fig. 1 have been drawn so that the condition (17) is satisfied.

But equation (17) is the condition which makes the linear vector function in (16) symmetric. This means that the vector set $\boldsymbol{a}$, $\boldsymbol{b}$ and the set $\boldsymbol{a}'$, $\boldsymbol{b}'$ can be interchanged, so the linear vector function (16) can also be written in the form:

$$\frac{\partial \boldsymbol{v}}{\partial t} = -(\boldsymbol{a}'\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{b}) \cdot \boldsymbol{r}. \tag{18}$$

It is found, by applying curl-differentiation, that the symmetric vector function (16, 18) is irrotational. Therefore, it can be represented as the ascendant of a potential ϕ:

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\nabla \phi. \tag{19}$$

It is easy to see that the potential is given by the expression

$$2\phi = \boldsymbol{r} \cdot (\boldsymbol{aa}' + \boldsymbol{bb}') \cdot \boldsymbol{r}\,. \tag{20}$$

∇-differentiation of (20) gives the sum of (16) and (18). The expression on the right in (20) is a homogeneous quadratic in $\boldsymbol{r}$, and the lines of constant ϕ are therefore a family of coaxial conics. The resultant buoyancy is the gradient to these conics.

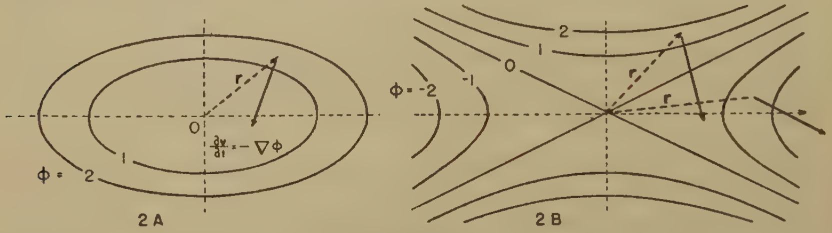


Figure 2.

The stability criterion can now be formulated very simply as follows. We have stability for displacements in every direction when the conics are ellipses; see Fig. 2a. And we have instability for displacements in the sectors of negative potential when the conics are hyperbolas; see Fig. 2b. The principal axes of the conics are the directions of maximum and minimum stability. It remains to find the orientation and sign of the principal axes for given values of the four vectors $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{b}$, $\boldsymbol{b}'$.

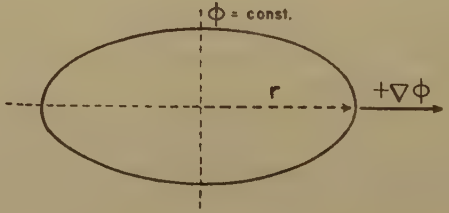


Figure 3.

A principal axis is directed along a radius vector which is perpendicular to the conic; see Fig. 3. It is therefore defined by the equation

$$n\boldsymbol{r} = \nabla\phi = (\boldsymbol{aa}' + \boldsymbol{bb}') \cdot \boldsymbol{r}\,, \tag{21}$$

where n is a scalar factor of proportionality, characteristic of the axis. A positive value of n signifies an axis of stability; a negative n means an axis of instability.

We assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are not parallel. Let $\boldsymbol{c}$ be a zonal unit vector, and let $\boldsymbol{a}^*$, $\boldsymbol{b}^*$, $\boldsymbol{c}$ be the reciprocal set of vectors to the set $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$. Then, according to a well known theorem, we have

$$\boldsymbol{r} = (\boldsymbol{a}\boldsymbol{a}^* + \boldsymbol{b}\boldsymbol{b}^*) \cdot \boldsymbol{r} \,. \tag{22}$$

Multiplication of (22) by n, and subsequent subtraction of (21), gives

$$0 = \{\boldsymbol{a}\,(n\boldsymbol{a}^* - \boldsymbol{a}') + \boldsymbol{b}\,(n\boldsymbol{b}^* - \boldsymbol{b}')\} \cdot \boldsymbol{r} \,.$$

Since $\boldsymbol{a}$ and $\boldsymbol{b}$ are assumed nonparallel, the two vectors $(n\boldsymbol{a}^* - \boldsymbol{a}')$, $(n\boldsymbol{b}^* - \boldsymbol{b}')$ are both perpendicular to $\boldsymbol{r}$, and therefore parallel. Thus:

$$(n\boldsymbol{a}^* - \boldsymbol{a}') \times (n\boldsymbol{b}^* - \boldsymbol{b}') = \boldsymbol{0} \,.$$

After performing the vector multiplication and subsequent scalar multiplication with the vector $\boldsymbol{a} \times \boldsymbol{b}$, we find

$$n^2 - (\boldsymbol{a} \cdot \boldsymbol{a}' + \boldsymbol{b} \cdot \boldsymbol{b}')n + (\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{a}' \times \boldsymbol{b}') = 0 \,. \tag{23}$$

The two roots of this equation determine the two principal axes of the family of conics (20). A positive root means an axis of stability. Thus, if both roots are positive, the conics are ellipses, and the vortex is stable for all displacements in the corresponding region.

The criterion for stability (two positive roots) is from (23)

$$\begin{aligned} B &= n_1 + n_2 = \boldsymbol{a} \cdot \boldsymbol{a}' + \boldsymbol{b} \cdot \boldsymbol{b}' > 0 \,, \\ C &= n_1 n_2 = (\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{a}' \times \boldsymbol{b}') > 0 \,. \end{aligned} \tag{24}$$

This stability criterion was given in 1939 by Høiland [2a], who, in the same paper, also gave a clear interpretation of the physical significance of the criterion.

The case which has been primarily discussed earlier is the indifferent case where one of the roots is zero, *i. e.*, $C = O$. This happens in the following three cases:

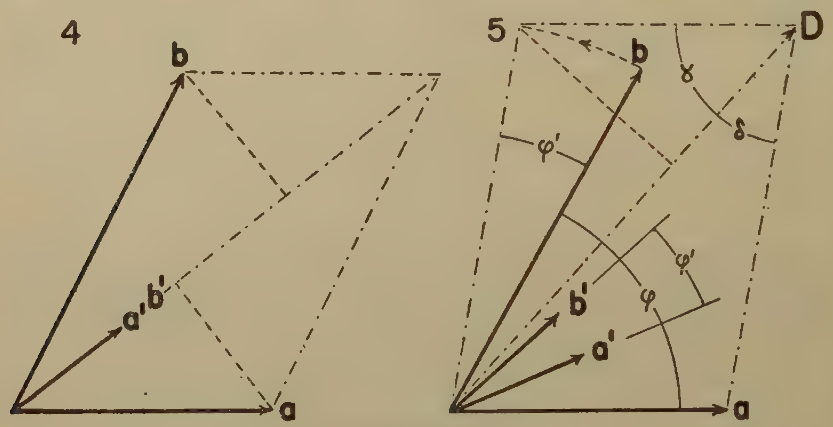
(a) $\boldsymbol{a} = \boldsymbol{0}$, the vortex is irrotational, and the centrifugal buoyancy disappears.

(b) $\boldsymbol{b} = \boldsymbol{0}$, the vortex is isentropic, and the gravitational buoyancy disappears.

(c) $\boldsymbol{a}' \times \boldsymbol{b}' = \boldsymbol{0}$. The vorticity surfaces coincide with the isentropic surfaces. This is equivalent to the statement that the wind shear measured along the isentropic surface is equal and opposite to the Coriolis parameter.

The indifferent case (c) is shown schematically in Fig. 4. It follows from the condition of balance (17) that, if the vectors $\boldsymbol{a}'$ and $\boldsymbol{b}'$ are parallel, they must lie along the diagonal of the parallelogram of the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The axis of indifference lies along the coinciding isentropic and vorticity lines. Obviously, displacements in the direction of these lines give rise to neither gravitational nor centrifugal buoyancy. The axis of maximum stability coincides with the diagonal, and from (24) the root of maximum stability is equal to the length of this diagonal.

In order to find the orientation of the axes of maximum and minimum stability in the general case, we must find the roots of the quadratic equation (23). With the abbreviated notations of (24), equation (23) becomes



Figures 4 and 5.

$$(25) \qquad n^2 - Bn + C = 0\,.$$

The solution of this equation is

$$(26) \qquad 2n = B \pm \sqrt{B^2 - 4C}\,,$$

where, from (24)

$$B^2 = (\boldsymbol{a} \cdot \boldsymbol{a}')^2 + (\boldsymbol{b} \cdot \boldsymbol{b}')^2 + 2\,(\boldsymbol{a} \cdot \boldsymbol{a}')\,(\boldsymbol{b} \cdot \boldsymbol{b}')\,.$$

We substitute for the first and second term on the right the following equivalent expressions (Lagrangean identities):

$$\begin{aligned}(\boldsymbol{a} \cdot \boldsymbol{a}')^2 &= a^2 - (\boldsymbol{a} \times \boldsymbol{a}')^2\,,\\ (\boldsymbol{b} \cdot \boldsymbol{b}')^2 &= b^2 - (\boldsymbol{b} \times \boldsymbol{b}')^2\,,\end{aligned}$$

and subsequently we reduce with the aid of the square of the expression in (17). We then find

$$B^2 = a^2 + b^2 + 2\,(\boldsymbol{a} \times \boldsymbol{a}') \cdot (\boldsymbol{b} \times \boldsymbol{b}') + 2\,(\boldsymbol{a} \cdot \boldsymbol{a}')\,(\boldsymbol{b} \cdot \boldsymbol{b}')\,.$$

Again from (24) we have

$$4C = 2\,(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{a}' \times \boldsymbol{b}') + 2\,(\boldsymbol{a} \cdot \boldsymbol{a}')\,(\boldsymbol{b} \cdot \boldsymbol{b}') - 2\,(\boldsymbol{a} \cdot \boldsymbol{b}')\,(\boldsymbol{a}' \cdot \boldsymbol{b})\,.$$

Combining these expressions for B^2 and $4C$ we find:

$$(27)\quad B^2 - 4C = a^2 + b^2 + 2\,(\boldsymbol{a} \cdot \boldsymbol{b})\,(\boldsymbol{a}' \cdot \boldsymbol{b}') - 2\,(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{a}' \times \boldsymbol{b}').$$

This expression is further simplified when the angles

$$\boldsymbol{a} \rightarrow \boldsymbol{b} = \varphi\,, \quad \boldsymbol{a}' \rightarrow \boldsymbol{b}' = \varphi'$$

are introduced; see Fig. 5. We then finally have

$$(28)\qquad B^2 - 4C = a^2 + b^2 + 2\,ab\cos(\varphi + \varphi') = D\,.$$

The discriminant, $B^2 - 4C$ in (26), is accordingly equal to the square of the magnitude of the diagonal $\boldsymbol{D}$ in a parallelogram whose sides are a and b, and whose side angle is $\varphi + \varphi'$. It is shown in Fig. 5 how this diagonal is found by a simple construction.

Introducing (28) in (26), the solution of the quadratic equation becomes

$$2n = B \pm D\,,$$

or

$$(29)\qquad B = n_1 + n_2\,; \quad D = n_1 - n_2\,.$$

The diagonal in the parallelogram in Fig. 5 is accordingly equal to the difference between the two roots. The diagonal is also simply related to the magnitudes of the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by the expression

$$(30)\qquad D = a\cos\gamma + b\cos\delta\,.$$

This relation will be used below to find the orientation of the principal axes.

The principal axes are found by introducing the two roots n_1 and n_2 in the equation (21). Let $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ denote the displacements along the two axes. We then have

$$(31)\qquad \begin{aligned} n_1\boldsymbol{r}_1 &= (\boldsymbol{a}\boldsymbol{a}' + \boldsymbol{b}\boldsymbol{b}') \cdot \boldsymbol{r}_1\,, \\ n_2\boldsymbol{r}_2 &= (\boldsymbol{a}\boldsymbol{a}' + \boldsymbol{b}\boldsymbol{b}') \cdot \boldsymbol{r}_2\,. \end{aligned}$$

If scalar multiplication is performed in the first equation by $\boldsymbol{r}_2$, and in the second equation by $\boldsymbol{r}_1$, the right-hand sides of both equations are the same, and we find by subtraction

$$(n_1 - n_2)\,\boldsymbol{r}_1 \cdot \boldsymbol{r}_2 = 0\,.$$

This equation gives the well known result that the principal axes are orthogonal when the two roots are different.

The equation for the conics, (20), becomes normalized in a system of co-ordinates whose axes coincide with the principal axes. In this system we have from (31)

$$\begin{aligned} n_1\boldsymbol{i} &= (\boldsymbol{aa}' + \boldsymbol{bb}')\cdot\boldsymbol{i}\,, \\ n_2\boldsymbol{j} &= (\boldsymbol{aa}' + \boldsymbol{bb}')\cdot\boldsymbol{j}\,. \end{aligned} \tag{32}$$

The magnitudes of these equations are

$$\begin{aligned} n_1 &= a_x a_x' + b_x b_x' = a\cos\alpha\cos\alpha' + b\cos\beta\cos\beta'\,, \\ n_2 &= a_y a_y' + b_y b_y' = a\sin\alpha\sin\alpha' + b\sin\beta\sin\beta'\,, \end{aligned} \tag{33}$$

where α, α', β, β' denote the angles between the x-axis and the four vectors $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{b}$, $\boldsymbol{b}'$; see Fig. 6. The difference between the two roots is from (29) and (33);

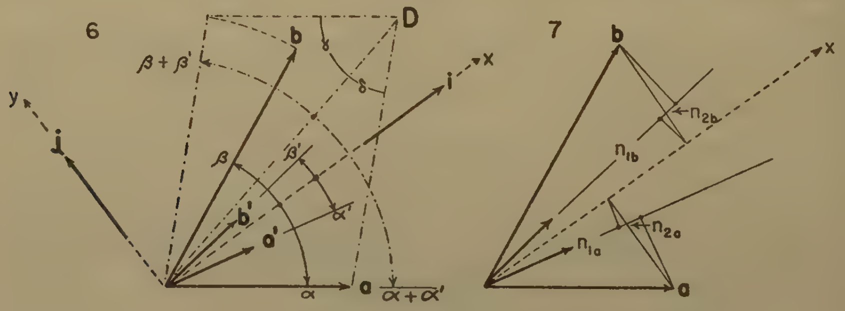
$$D = n_1 - n_2 = a\cos(\alpha + \alpha') + b\cos(\beta + \beta')\,. \tag{34}$$

This expression must be identical to (30). It follows that the principal axes have such an orientation that

$$\alpha + \alpha' = \gamma\,; \quad \beta + \beta' = \delta\,.$$

Therefore it appears from Fig. 6 that the principal axis of maximum stability (the x-axis) bisects the angle between the diagonal $\boldsymbol{D}$ and the vector $\boldsymbol{a}'$.

The magnitude of the two roots can also be found graphically on the same diagram, as shown in Fig. 7. According to (33) each root is the sum of two terms. One of these terms depends upon the centrifugal buoyancy and one term depends upon the gravitational buoyancy. The centrifugal term in the root of maximum stability is $n_{1a} = a\cos\alpha\cos\alpha'$. This term is found graphically by projecting the x-component



Figures 6 and 7.

of $\boldsymbol{a}$ on the $\boldsymbol{a}'$-axis. The gravitational term of the same root, n_{1b}, is found by performing the same construction with reference to the b-vectors. The centrifugal term of the root of minimum stability, $n_{2a} = a \sin \alpha \sin \alpha'$, is the additional segment on the $\boldsymbol{a}'$-axis, when the vector $\boldsymbol{a}$ is projected directly on this axis. The gravitational term of this root, n_{2b}, is the corresponding segment on the $\boldsymbol{b}'$-axis.

The diagram in Figs. 6 and 7 contains in graphical form all the required information on the state of stability of the vortex.

The normalized form of the equation for the family of conics is obtained by introducing the conditions (32) in the general equation (20). Writing here $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$, and reducing with the aid of (32), we find

$$2\phi = n_1 x^2 + n_2 y^2 . \tag{35}$$

Substituting this value of ϕ in (19), the perturbation equation becomes

$$\frac{\partial \boldsymbol{v}}{\partial t} = - n_1 x \boldsymbol{i} - n_2 y \boldsymbol{j} . \tag{36}$$

From this equation we can readily find the "free" oscillations which the particles (zonal rings) would perform under the joint action of centrifugal and gravitational buoyancy. It is found that the motion is the resultant of two harmonic oscillations, one along the axis of maximum stability with the frequency $\sqrt{n_1}$ and one along the axis of minimum stability with the frequency $\sqrt{n_2}$. However, it is of little interest to study this motion, since the actual motion is governed by the kinematic constraint at the boundaries. The actual motion can be obtained only by including the effect of local pressure changes in the perturbation equation and using the complete set of equations and boundary conditions.

The analysis which has been presented above for a zonal current applies as well to geostrophic currents of arbitrary orientation. The vectors $\boldsymbol{b}$, $\boldsymbol{b}'$, which define the gravitational buoyancy, has exactly the same meaning as in the zonal case. The vectors $\boldsymbol{a}$, $\boldsymbol{a}'$ have the same physical meaning but a somewhat different form, as they are expressed in terms of the wind relative to the earth.

SUMMARY

The dynamic stability of zonal currents is controlled by the joint action of centrifugal and gravitational buoyancy. The resultant of these two buoyancies is a symmetric linear vector function of the meridional displacement. With the aid of the theory of linear vector functions, the directions of maximum and minimum stability and the

magnitudes of these stabilities are obtained. The results are found graphically by a convenient elementary construction.

REFERENCES

[1.] Solberg, Halvor

1936. Le mouvement d'inertie de l'atmosphére stable et son rôle dans la théorie des cyclones. P. V. Mét. Un. géod. géophys. int. Mem. Disc., *2:* 66–82.

[2.] Høiland, Einar

1939. On the interpretation and application of the circulation theorems of V. Bjerknes. Arch. Math. Naturv., Bd. 42, No. 5: 1–69.

1941. On the stability of the circular vortex. Avh. norske VidenskAkad., Mat.-naturv. Kl., No. 11: 1–24.

[3.] Kleinschmidt, Ernst

1941. Stabilitätstheorie des geostrophischen Windfeldes. Ann. Hydrogr. mar. Met., *69:* 305–325.

1941. Zur Theorie der labilen Anordnung. Met. Z., *58:* 157–163.

[4.] Van Mieghem, Jacques

1945. Forme intrinsèque du critère d'instabilité dynamique de E. Kleinschmidt. Bull. Acad. Belg. Cl. Sci., (5) *30:* 19–33.

[5.] Fjørtoft, Ragnar

1946. On the frontogenesis and cycogenesis in the atmosphere. Part 1. On the stability of the stationary circular vortex. Geofys. Publ., *16* (5): 1–28.

[6.] Bjerknes, Vilhelm

1937. Application of line integral theorems to the hydrodynamics of terrestrial and cosmic vortices. Astrophys. Norwegica, *2:* 263–339.

ON DISPLACEMENTS AND INTENSITY CHANGES OF ATMOSPHERIC VORTICES

By
C.-G. ROSSBY

Institutet för Meteorologi
Stockholm, Sweden

Aerological investigations conducted during the last few years have revealed that the long, meandering disturbances characteristic of the middle latitude belt of westerlies in the upper troposphere frequently increase in amplitude until troughs and ridges finally are cut off from the main current, thus forming cold cyclonic vortices on the equatorial side of the west-wind belt and warm anticyclonic vortices on the north side thereof. True vortices, with nearly vertical axes, are thus characteristic phenomena not only of equatorial regions where they appear as tropical cyclones but also of middle latitudes. The middle latitude vortices seem to be phenomena of great stability which have sometimes been observed to persist for several weeks.

In view of these observations it may be of interest to call attention to a dynamic factor which appears to be of some significance to the intensity fluctuations and displacements of simple vortices with vertical axes. A few elementary statements with regard to this factor will be made below and an attempt will be made to suggest a few practical consequences of the results obtained.

We shall consider first a simple vortex, in which the air *at a given instant* spins around the vertical axis with a relative angular velocity ω, which may be a function of r, the horizontal distance from the axis, and of z, the height above sea level, but *not* of the azimuth angle φ. Outside a certain distance R, the outer boundary of the vortex, the air does not participate in the vortex motion.

In such a vortex the Coriolis' force acting on an air parcel to the north of the center will be somewhat greater than the Coriolis' force acting on a correspondingly located parcel south of the center, the difference depending upon the rate of variation of the Coriolis' parameter with latitude. Thus the vortex will be acted upon by a resultant force F, the magnitude of which may be computed quite readily. If one introduces a coordinate system with its center in the axis of the vortex, with the x-axis pointing towards the east and the y-axis towards the north, it follows that the distribution of the Coriolis' parameter (f) may be given by

$$f = f_0 + \beta y = f_0 + \beta r \sin\varphi \, , \tag{1}$$

where f_0 is the value of this parameter in the latitude of the center of the vortex and β its rate of change northward. Variations in β will be neglected. The resultant F of the Coriolis' forces acting on a horizontal unit slice is directed northward and has the value

$$F = \rho \int_0^R r dr \int_0^{2\pi} f r \omega \sin \varphi \, d\varphi \, . \tag{2}$$

The relative angular velocity ω is positive for cyclonic motion. Substitution of (1) in (2) gives

$$F = \beta \rho \pi \int_0^R r^3 \omega dr \, . \tag{3}$$

If $\bar{\omega}$ represents a mean value of the angular velocity within the vortex, one finds

$$F = \beta M \frac{\bar{\omega} R^2}{4} \, , \tag{4}$$

where

$$M = \rho \pi R^2 \tag{5}$$

represents the mass of a horizontal unit slice of the vortex. Since there are no pressure gradients in the environment to balance the force F, it is evident that the vortex will receive an acceleration, northward for $\bar{\omega} > 0$ (cyclonic vortices) and southward for $\bar{\omega} < 0$ (anticyclonic vortices).[1]

Before computing the resulting displacement we shall make a comparison of the kinetic energy of the vortex in two different latitude positions. In this comparison we shall assume the vortex to be completely barotropic, so that the velocity distribution is independent of height. Under those circumstances the distribution of mass will remain practically constant and a fluid ring in a given distance r from the axis will therefore remain in approximately the same distance throughout the life history of the vortex. The absolute circulation C of such a fluid ring remains constant, in spite of the displacement of the vortex. If one indicates the mean value of the Coriolis' parameter in the first position as f_1 and in the second position as f_2, it follows that

---

[1] The acceleration discussed above should not be confused with the so-called rotor effect, which depends upon the presence of a circulation around the vortex in the outer fluid and upon a translation of the vortex relative to the surrounding medium.

$$2\pi r^2\left(\omega_1+\frac{f_1}{2}\right)=2\pi r^2\left(\omega_2+\frac{f_2}{2}\right), \tag{6}$$

or, in differential form,

$$\delta\omega=-\tfrac{1}{2}\,\delta f\,. \tag{7}$$

The kinetic energy K of a horizontal slice of unit thickness is given by

$$K=\pi\rho\int_0^R r^3\omega^2 dr\,. \tag{8}$$

For the case of solid rotation one finds

$$K=\pi\rho\omega^2\frac{R^4}{4}\,. \tag{9}$$

Since in this case the vertical component of the absolute vorticity Z is given by

$$Z=2\omega+f\,, \tag{10}$$

and since Z, according to (6), must remain constant during the displacement of the vortex, it follows that the kinetic energy conveniently may be written in the form

$$K=\frac{\pi}{16}\,\rho\,(Z-f)^2R^4=\frac{M}{16}\,(Z-f)^2R^2\,, \tag{11}$$

showing that solid cyclonic barotropic vortices ($Z>f$) suffer a decrease in kinetic energy with northward displacements but gain kinetic energy if forced to move southward.

If the vortex possesses a translational velocity v northward through the surrounding medium, and if one neglects the kinetic energy of the environment, its energy equation must be of the form

$$K+M\frac{v^2}{2}=\text{constant}\,. \tag{12}$$

Individual differentiation with time gives

$$\frac{\partial K}{\partial f}\,\frac{df}{dt}+Mv\frac{dv}{dt}=0\,, \tag{13}$$

or, since

$$\frac{df}{dt}=\beta v\,, \tag{14}$$

$$M\frac{dv}{dt}=-\beta\frac{\partial K}{\partial f}\,. \tag{15}$$

It follows from (11) that

$$-\beta \frac{\partial K}{\partial f} = \frac{\pi}{8} \rho (Z - f) \beta R^4 = M \frac{\beta \omega R^2}{4}. \tag{16}$$

Thus

$$M \frac{dv}{dt} = M \frac{\beta \omega R^2}{4}, \tag{17}$$

in complete agreement with the expression for the resultant Coriolis' force given in (4).

To compute the magnitude of the translational velocity produced by this force, we might make use of (12), assuming that $v = 0$ in the first position ($f = f_1$). It follows from (11) and (12) that

$$\frac{v_2^2}{2} = \frac{R^2}{16} [(Z - f_1)^2 - (Z - f_2)^2], \tag{18}$$

or

$$v_2^2 = \frac{R^2}{4} \left[\left(Z - \frac{f_1 + f_2}{2} \right) (f_2 - f_1) \right]. \tag{19}$$

To arrive at the significance of this result we shall apply it to a small, intense vortex of hurricane type. It is then permissible to neglect f in comparison with Z, and (19) reduces to

$$v_2 = \frac{R}{2} \sqrt{Z(f_2 - f_1)} = \sqrt{\frac{R^2 \omega (f_2 - f_1)}{2}}. \tag{20}$$

Assuming $R\omega = 25$ mps, $R = 200$ km, $f_2 - f_1 = 10^{-5}$ sec^{-1}, one finds

$$v_2 = 5 \text{ mps}$$

for the northward drift. This value would be reduced, perhaps by fifty per cent or more, if one were to include in (12) also the kinetic energy of the environment. The order of magnitude of the result appears to be in good agreement with the tendency of tropical hurricanes to drift somewhat to the north and right of the trade wind current in which they are embedded.

We shall next attempt to determine the resulting Coriolis' force on a particularly simple barocline and anticyclonic vortex. For this purpose we select a cold core anticyclonic vortex with the air density ρ, surrounded by a lighter air mass of density ρ', at rest relative to the surface of the earth. To simplify the numerical work as much as possible, we shall select a vortex in which the absolute vorticity vanishes, so that

(21) $$f + \frac{1}{r}\frac{\partial ur}{\partial r} = 0\,,$$

u being the tangential linear velocity, counted positive for cyclonic motion, and f the mean (or center) value of the Coriolis' parameter. It follows that

(22) $$u = -\frac{fr}{2}\,.$$

This represents an anticyclonic vortex with a cold dome of maximum steepness. To determine the shape of the dome one may make use of the gradient wind equation, which in this case takes the form

(23) $$\frac{u^2}{r} + fu - \gamma\frac{\partial D}{\partial r} = 0\,,$$

γ being an abbreviation for $\frac{\rho-\rho'}{\rho}\,g$ and D being the depth of a cold air column at a given distance r from the center. From a combination of (22) and (23) one finds

(24) $$\frac{\partial D}{\partial r} = -\frac{f^2 r}{4\gamma}\,,$$

and, upon integration,

(25) $$D = \frac{f^2}{8\gamma}(R^2 - r^2)\,,$$

R being the distance in which the cold dome boundary reaches the ground. If V is the total volume of the cold air, one finds

(26) $$fR^2 = 4\sqrt{\frac{\gamma V}{\pi}}\,,$$

and hence

(27) $$D = \frac{2V}{\pi R^4}(R^2 - r^2)\,.$$

The kinetic energy KE of the vortex is readily computed. One finds

(28) $$KE = \frac{\rho V}{24}f^2R^2 = \frac{\rho V}{6}\sqrt{\frac{\gamma V}{\pi}}\,f\,.$$

The variable part of the potential energy of the system (PE) is given by

(29) $$PE = \int_0^R \rho\gamma\frac{D^2}{2}\,2\pi r dr = \frac{\rho V}{6}\sqrt{\frac{\gamma V}{\pi}}\,f$$

and thus potential and kinetic energy are equal. During any displacement of this vortex the total volume V or mass M must remain constant. The total energy E is given by

$$E = \frac{\rho V}{3}\sqrt{\frac{\gamma V}{\pi}}\,f\,, \tag{30}$$

and this quantity is thus proportional to the Coriolis' parameter f. The individual variation of E is given by

$$\frac{dE}{dt} = \frac{\rho V}{3}\sqrt{\frac{\gamma V}{\pi}}\,\beta v = \frac{\mathrm{M}}{3}\sqrt{\frac{\gamma V}{\pi}}\,\beta v\,, \tag{31}$$

v being the rate of translation northward.

We shall next compute directly the force F acting upon this vortex as a result of the nonbalance of the Coriolis' forces. One finds, as before, if F is counted positive northward,

$$F = \rho \int_0^R rDdr \int_0^{2\pi} (f_0 + \beta r \sin\varphi)\, u \sin\varphi d\varphi\,, \tag{32}$$

and, after proper substitutions,

$$F = -M\frac{\beta}{3}\sqrt{\frac{\gamma V}{\pi}}\,. \tag{33}$$

If one neglects the kinetic and potential energy of the environment, the energy equation for this system must be of the form

$$M\frac{v^2}{2} + E = \text{constant}\,, \tag{34}$$

or

$$Mv\frac{dv}{dt} = -\frac{dE}{dt} = -\frac{M}{3}\sqrt{\frac{\gamma V}{\pi}}\,\beta v\,. \tag{35}$$

It follows that

$$M\frac{dv}{dt} = -M\frac{\beta}{3}\sqrt{\frac{\gamma V}{\pi}}\,, \tag{36}$$

and this expression is identical with the resultant force F given in (33).

From the two simple examples treated so far we may infer that, if the sense of rotation remains the same at all levels:

A. Axially symmetric cyclonic vortices are driven northward and the energy associated with the rotational motion is decreased and changed into a corresponding increase of the translational kinetic energy;

B. Axially symmetric anticyclonic vortices are driven southward and *the potential and kinetic energy associated with the rotational motion is decreased and changed into a corresponding increase of the translational kinetic energy.*

One should expect that the translation of vortices northward or southward would give rise to Coriolis' forces of the form Mfv, and that these forces would deflect northward moving (cyclonic) vortices towards the east and southward moving vortices towards the west. However, in an earlier study [1] the author has shown that in a nondivergent horizontal fluid sheet of density ρ' and subjected to a constant Coriolis' parameter f, a vertical cylinder of radius a, moving through the fluid at a constant speed v, will be acted upon by a resultant pressure gradient force of the magnitude $\rho'\pi a^2 fv$ which is directed 90° to the left of the translation.[2] This pressure gradient balances the Coriolis' force of the translation completely, or very nearly so, and will thus permit the vortex to move practically straight northward or southward in response to the resultant force F. In the case of a cyclonic vortex of hurricane type which is carried in an easterly trade wind current, the effect of the force F would be to deflect the direction of motion of the storm somewhat poleward of the direction of the steering current.

The resultant force F stems from the local lack of balance between the horizontal pressure gradient on the one hand and the sum of Coriolis' and centrifugal forces on the other hand. One may ask if this local unbalance might not produce form changes whereby the vortex would lose its initial symmetry and ultimately acquire a shape characterized by dynamic balance in all points. If such equilibrium shapes exist, as suggested by Bjerknes and Holmboe [2], they must be characterized by the condition that F vanishes.

It has not yet been possible to undertake an exhaustive study of such possible equilibrium shapes, but a preliminaty analysis indicates that the simple concentrated vortices considered in the present paper are far removed from internal equilibrium. In the case of purely

---

[2] This result is readily obtained if one makes use of the Bernoulli equation. If the observed motion is steady relative to a coördinate system travelling with the speed c in the positive x-direction (which may have an arbitrary direction as long as f is considered constant), this equation takes the form

$$\frac{d}{dt}\left[p + \rho' \frac{(U-c)^2 + V^2}{2} + \rho' f c y\right] = 0,$$

U and V being the x- and y-components of motion. The contributions to the pressure resultant from the quadratic kinetic energy term cancel out as long as the motion in the environment remains symmetric with respect to the x-axis, and this condition is satisfied if the motion is nondivergent and if f is considered as a constant.

horizontal, nondivergent motion, the problem may be approached as follows:

The horizontal velocity components may be expressed through a stream function ψ, and are given by

$$V = \frac{\partial \psi}{\partial x}, \quad U = -\frac{\partial \psi}{\partial y}. \tag{37}$$

The equations of motion may then be written:

$$\rho \frac{dU}{dt} = \rho f \frac{\partial \psi}{\partial x} - \frac{\partial p}{\partial x}, \tag{38}$$

$$\rho \frac{dV}{dt} = \rho f \frac{\partial \psi}{\partial y} - \frac{\partial p}{\partial y}, \tag{39}$$

or, if the y-axis is directed northward

$$\rho \frac{dU}{dt} = \frac{\partial \Phi}{\partial x}, \tag{40}$$

$$\rho \frac{dV}{dt} = \frac{\partial \Phi}{\partial y} - \rho \beta \psi, \tag{41}$$

where

$$\Phi = \rho f \psi - p. \tag{42}$$

If the atmosphere surrounding the vortex is to be at rest, it follows that ψ must approach zero in great distances from the center of the vortex and p must approach a constant value. If the circulation along the various stream lines surrounding the vortex center is always of the same sense, it follows that ψ must vary in a monotone fashion from the center and outward. Thus, in case of a cyclonic vortex, ψ must grow from a negative minimum value to zero, while in the case of an anticyclonic vortex, ψ must fall from a positive maximum value to zero.

If one integrates (41) over the entire area of such a monotone vortex, it follows that

$$F = \iint \frac{dV}{dt} \rho \, dxdy = -\rho\beta \iint \psi \, dxdy, \tag{43}$$

and this expression is positive for arbitrary cyclonic monotone vortices and negative for arbitrary anticyclonic monotone vortices. It is easily shown that (43), for the case of axially symmetric vortices, leads to the same value for F as was given in (3). Within the limitations set by the simplified theory, one must conclude that there are no monotone vortices of equilibrium type; cyclonic vortices are always subjected to a resultant force directed northward, while in the case of anticyclonic vortices the force is directed southward.

The results of this analysis remain unchanged if one superimposes upon the entire system a zonal current U_0, which may depend upon y. To the previous value of the stream function ψ one must then add a basic function

$$\psi_0 = -\int U_0 dy ,$$

and to the previous pressure p a basic pressure field

$$p_0 = -\rho\int\int U_0 dy ,$$

corresponding to the geostrophic requirement.

The effect of axially symmetric flow out from or in towards the center of the vortex is easily evaluated by an analogous method. Such motions may be described with the aid of a velocity potential θ, which is a function of r only. The corresponding velocity components would be

$$(46) \qquad U = \frac{\partial\theta}{\partial x} = \frac{x}{r}\frac{\partial\theta}{\partial r}, \quad V = \frac{\partial\theta}{\partial y} = \frac{y}{r}\frac{\partial\theta}{\partial r} .$$

The horizontal divergence would be given by

$$(47) \qquad \operatorname{div}_2 \mathbf{V} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\theta}{\partial r}\right),$$

and to insure that the area integral of this quantity vanish, one must assume that $\frac{\partial\theta}{\partial r}$, for large values of r, decreases more rapidly than r^{-1}. Radial outflow would then correspond to a divergence area at the center, surrounded by a ring-shaped convergence area. In the case of radial inflow, the distribution of convergence and divergence areas would be reversed. Radial outflow would correspond to a θ-function increasing from a negative minimum at the center to zero at the periphery. Inflow would correspond to a θ-function decreasing from a positive maximum at the center to zero at the periphery. One then finds that these radial motions give rise to a force H, directed eastward and having the value

$$(48) \qquad H = -\rho\beta\int\int\theta \, dx \, dy .$$

Thus vortices with radial outflow are accelerated eastward, vortices with radial inflow are accelerated westward.

Displacements eastward or westward of barocline vortices are not accompanied by immediate and direct changes in the thermal wind balance, as is the case when such vortices are displaced northward or southward. Therefore, we shall devote the remainder of this paper to a discussion of the effects of the force F.

The vortices analyzed above are unstable in the sense that they must undergo a continuous displacement and decrease in energy until their rotational energy has been converted into translational energy. It is clear, however, that the entire process is most appropriately described in terms of a relaxation time, defining the total length of the life cycle of the vortex.

In the analysis of the displacement of a cold core vortex presented above, we neglected to consider the kinetic energy of the environment. Inspection of (26) shows that the sea level radius R of the cold core must vary in such a manner that

$$R \approx \frac{1}{\sqrt{f}},$$

and this result, in combination with (26), shows that

$$D_{\max} \approx f .$$

It follows that a displacement of the vortex from 45° N to 30° N would result in an increase in R of about 20% and a decrease in the maximum height of the dome to about 71% of its initial value. This divergence of the cold air must result in convergence of the lighter air aloft and the generation of cyclonic circulation at these levels. As a consequence, the total force F, in this case directed southward, will be diminished, and less rotational (potential and kinetic) energy becomes available for conversion into translational energy.

In the case of purely barotropic vortices we have shown that the displacement in latitude results in a fairly rapid dispersion of the rotational kinetic energy. The qualitative analysis of the effect of low level divergence presented above shows that this dispersion may be reduced, and to some extent replaced, by a concentration of rotational kinetic energy aloft. To explore how far this concentration may proceed it is useful to inspect anew the structure of the force of displacement F, as given in (3). It is evident from this expression, in which ω is multiplied by r^3, that in a cold core vortex the negative contribution to F from a feeble but widespread anticyclonic circulation in the lower layers may dominate over the positive contribution to F from a more intense but restricted cyclonic circulation aloft[3]. *The*

---

[3] It follows from simple dimensional considerations that a maximum low level anticyclonic wind of 4 mps at a distance of 600 km from the axis would be sufficient to balance a maximum high level cyclonic wind of 64 mps at 200 km from the center, in so far as the force F is concerned.

complex barocline vortex with a broad anticyclonic base may thus continue its displacement southward in spite of strong cyclonic winds aloft until the resultant value of F *vanishes. The vortex has then reached its equilibrium latitude and cyclonic vorticity has been brought southward at high levels.*

Since the source of the rotational cyclonic energy aloft must come from the energy supply in the lower layers, it is important to show that the intensification of the circulation aloft is compensated for by a decrease in the potential energy of the system, *i. e.*, through a sinking of the cold air. Such a transfer of energy upward is generally not possible as long as the vortex remains in a fixed latitude, because of the normally prevailing inertia stability. It would appear from the preceding analysis, however, that a cold core vortex, anticyclonic in the lower layers but cyclonic aloft and subjected to a slight decrease in f, must experience an initial centripetal acceleration in the upper layers and an initial centrifugal acceleration below. Through the resulting direct circulation around the solenoids, with descending motion near the center and ascending motion in the periphery, part of the potential energy will be converted into cyclonic kinetic energy aloft. In this connection it is perhaps of interest to mention that the cold upper level cyclones which form as a result of excessive meandering normally seem to develop over broad low-level anticyclones [3]. It is not suggested that the initial cutting-off process should be attributed to subsidence, since synoptic experience clearly indicates that the mechanism of cutting-off is associated with changes in the field of motion upstream, but it is suggested that the subsequent intensification aloft may be associated with subsidence in the underlying cold anticyclone as it is displaced southward.

In connection with some recent aerological investigations of flow patterns over the western Pacific (region of the Marianas), Riehl [4] finds indications that typhoons have a tendency to develop under anticyclonic cells which have been observed to exist at the 30,000–50,000-foot levels over this part of the Pacific. Riehl attributes the formation of typhoon circulations in the lower and middle troposphere to divergence associated with these upper level anticyclones. It is not improbable that centrifugal accelerations associated with the southward displacement of such upper anticyclones may account for the dynamic unbalance required to initiate the divergence postulated by Riehl. The subsequent, continued removal of air would presumably require a different and more effective mechanism. This problem, recently investigated by Sawyer [5], will not be considered here.

SUMMARY

It is shown that, in any horizontal plane in the atmosphere, isolated closed circulation cells of arbitrary shape are subjected to a resultant force F which, in the case of cyclonic circulations, is directed northward, and in the case of anticyclonic circulations southward. It follows that simple atmospheric vortices with vertical axes have a tendency to move across the latitude circles in such a direction that the absolute value of their circulation relative to the surface of the earth decreases. Furthermore, it is shown, with the aid of two simple examples, that the decrease in potential and rotational kinetic energy of a vortex so displaced is sufficient to account for the increase in translational kinetic energy produced by the force F. It is well known that under normal conditions a symmetric, anticyclonically rotating dome of cold air is prevented by inertia stability from spreading out while remaining in a fixed location. This stability is overcome by the force F, which brings about a displacement of the cold dome equatorwards into regions where the thermal wind balance is such as to permit further sinking and spreading out.

The southward displacement and consequent sinking of cold anticyclonic domes must result in the generation of cyclonic circulations at higher levels. Thus composite, cold-core vortices may arise with a broad anticyclonic base and a concentrated cyclonic circulation aloft, and these vortices must continue to move southward with increasing cyclonic circulation aloft, until the resultant value of F, obtained through integration over the entire vertical extent of the vortex, vanishes. The composite vortex has then reached its equilibrium latitude; in the course of its displacement it has contributed to the southward transport of cyclonic vorticity aloft and to its accumulation near the equilibrium latitude, presumably at the tropopause level.

In a similar manner warm, anticyclonically rotating air masses embedded between the troposphere and the stratosphere, must tend to drift southward and spread out laterally. In so doing they must create a certain amount of convergence and ascending air motion in the underlying parts of the troposphere. If the underlying portion of the atmosphere is convectively unstable, this convergence may facilitate the development of tropical storms.

REFERENCES

(1) Rossby, C.-G.
1936. Dynamics of steady ocean currents in the light of experimental fluid mechanics. Pap. Phys. Oceanogr. Meteorol., *5* (1): 1–43.

(2) Bjerknes, J. and J. Holmboe
1944. On the theory of cyclones. J. Meteorol., *1* (1): 1–22.

(3) Palmén, Erik
1948. Origin of high-level cyclones south of the maximum westerlies. U. G. G. I., Reunion d'Oslo. Progr. Rés. Mémoires, pp. 35–37.

(4) Riehl, Herbert
1948. On the formation of typhoons. Tech. Rep. (mimeo.) Dept. Meteorol., Univ. Chicago.

(5) Sawyer, J. S.
1947. Notes on the theory of tropical cyclones. Quart. J. R. met. Soc., *73* (315–316): 101–126.

THE DYNAMICS OF LARGE HORIZONTAL EDDIES (AXES VERTICAL) IN THE OCEAN OFF SOUTHERN CALIFORNIA[1]

By

GEORGE F. McEWEN[2]

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

A striking feature shown on charts of horizontal circulation of the ocean in the southern California region, Lat. 32° to 34.5° N., is a large counterclockwise eddy between a northward inshore current and a southward offshore current. Charts arranged in time order from spring, at intervals of about two weeks, show a decrease in velocity and an increase in size of the eddy. This suggests that these eddies may be successive stages in the decay of an initial state of maximum intensity and relatively small extent. In order to simplify the dynamical problem thus presented, equivalent circular eddies are derived by averaging the velocities of the actual oval-shaped ones.

Each of three series of such circular eddies conformed approximately to the principle of the conservation of angular momentum, and the kinetic energy was found to decrease according to a formula of the dissipation of energy caused by lateral eddy viscosity. Values of the coefficient of eddy viscosity computed by means of this formula ranged from 3×10^4 to 6×10^5 c. g. s.

A combination of the two general hydrodynamic differential equations of horizontal motion in the sea, involving the deflecting force due to the earth's rotation as well as pressure gradient and lateral turbulence, led to a single differential equation of the diffusion of vorticity. A solution of this equation was found corresponding to constant initial vorticity over a circular area of appropriate radius

---

[1] Presented at the meeting of the American Physical Society, University of California at Los Angeles, January 3, 1948.

Contributions from the Scripps Institution of Oceanography, New Series, No. 383.

[2] It is a pleasure to acknowledge the help of Hector Venancio Iglesias, Lt. Comdr. in the Argentine Navy, one of my students, for his skillful execution of various supplementary computations that involved methods of interpolation, differentiation and integration.

and zero vorticity outside of this area. Application of Stokes theorem (Lamb, 1945: 31–36) relating circulation to vorticity provides a means of computing the velocity corresponding to any time and distance from the center of the eddy. Substituting in this expression for velocity the value of the coefficient of eddy viscosity (found from the rate of energy dissipation) and a value of the initial radius (estimated from an inspection of the charts) yields the theoretical velocity except for a constant factor. This factor, which is also the maximum initial velocity, was determined by comparison of the values of the foregoing expression with the velocities at the earliest date of observations in each series. Thus the velocity-distance relation can be predicted for any later time. The general agreement of the mathematical model with observations materially reinforces the evidence in support of the thesis suggested by an inspection of current charts arranged according to time.

THE HORIZONTAL CIRCULATION OFF SOUTHERN CALIFORNIA INFERRED FROM DYNAMIC HEIGHT ANOMALIES ACCORDING TO THE BJERKNES METHOD

The charts of horizontal circulation (Figs. 1 to 10) off southern California show certain striking features typical of this region. They are a selection based upon systematic seasonal surveys (Sverdrup and Staff, 1938: charts on pp. 14, 16, 17, 19, 21; 1939: charts on pp. 68, 69, 71–75; 1940: MS. charts; 1941: charts on pp. 250–264) made in collaboration with the California State Fish and Game Commission. Attention is called especially to the large counterclockwise eddies, centered at about Lat. 33.5° N. and Long. 119.5° W., which form a striking part of the general circulation pattern of prevailing northward inshore currents and southward offshore currents. Although recognized by various oceanographers, these eddies have not been the subject of detailed investigation. An inspection of Figs. 1 to 5, showing the circulation at the 50-meter level during 1940, suggests a succession of stages of an eddy of maximum initial intensity and relatively small extent. This percept is also suggested by Figs. 6 to 10 showing the surface circulation during 1941, although the period of observation is about two months later and although the first eddy has a larger area than that of 1940. In order to simplify the problem and to emphasize the essential characteristics, equivalent circular eddies were obtained by averaging the velocity between each pair of streamlines and averaging the distances to the center of the actual oval-shaped eddy. Figs. 11 to 14 show such equivalent circular eddies at the surface for 1940, and Figs. 15 to 18, for the same year, show equiva-

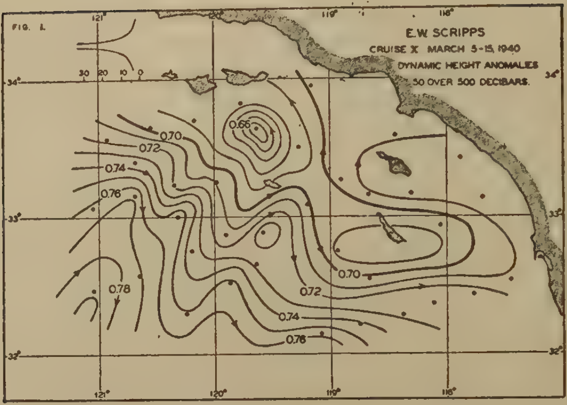


FIG. 1.
E.W. SCRIPPS
CRUISE X MARCH 5-15, 1940
DYNAMIC HEIGHT ANOMALIES
50 OVER 500 DECIBARS.
121°
120°
119°
118°
34°
33°
32°
30 20 10 0
0.66
0.70
0.72
0.74
0.76
0.78
0.70
0.72
0.74
0.76

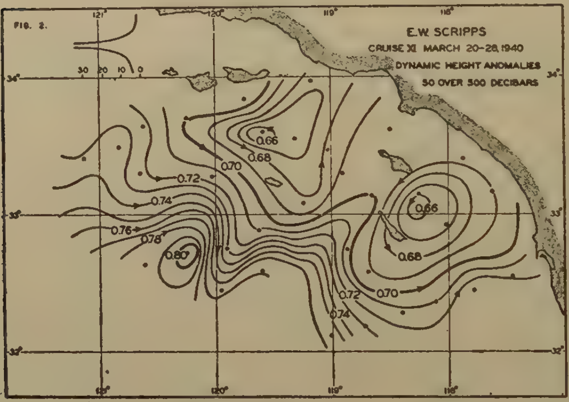


FIG. 2.
E.W. SCRIPPS
CRUISE XI MARCH 20-28, 1940
DYNAMIC HEIGHT ANOMALIES
50 OVER 500 DECIBARS
121°
120°
119°
118°
34°
33°
32°
30 20 10 0
0.66
0.68
0.70
0.72
0.74
0.76
0.78
0.80
0.66
0.68
0.70
0.72
0.74

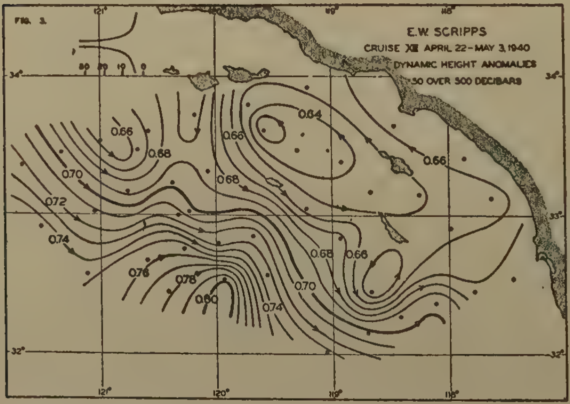


FIG. 3.
E.W. SCRIPPS
CRUISE XII APRIL 22-MAY 3, 1940
DYNAMIC HEIGHT ANOMALIES
50 OVER 500 DECIBARS
121°
120°
119°
118°
34°
33°
32°
30 20 10 0
0.64
0.66
0.66
0.68
0.66
0.68
0.70
0.72
0.74
0.76
0.78
0.80
0.74
0.70
0.68
0.66

Figures 1–3. Dynamic Height Anomalies (current boundaries) 50 over 500 Decibars, 1940 lent circular eddies at the 50-meter level. The average velocity between each pair of circles is inversely proportional to the distance between them, according to the Bjerknes theory. The simplest curves showing the relation of velocity to distance from the center, corresponding to each circular eddy, are shown by the heavy lines in the same figures. In general, the curves show that the velocity increases from zero at the center of the eddy to a maximum, and then decreases. The position of maximum velocity of the surface eddy, at the earliest date, is about 12 kilometers from the center and shifts farther from the center at later stages. The eddy at the 50-meter

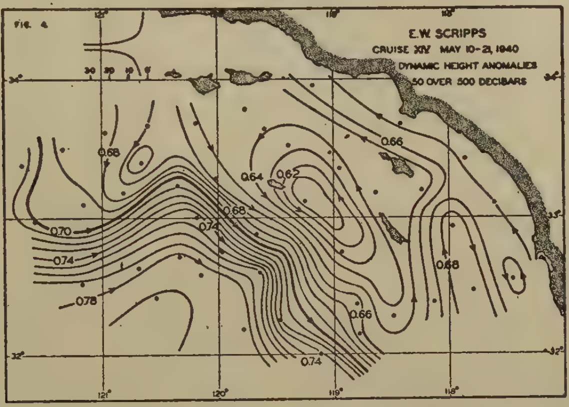


FIG. 4
E.W. SCRIPPS
CRUISE XIV MAY 10-21, 1940
DYNAMIC HEIGHT ANOMALIES
50 OVER 500 DECIBARS
121°
120°
119°
118°
34°
33°
32°
30
20
10
0.68
0.64
0.62
0.66
0.70
0.74
0.78

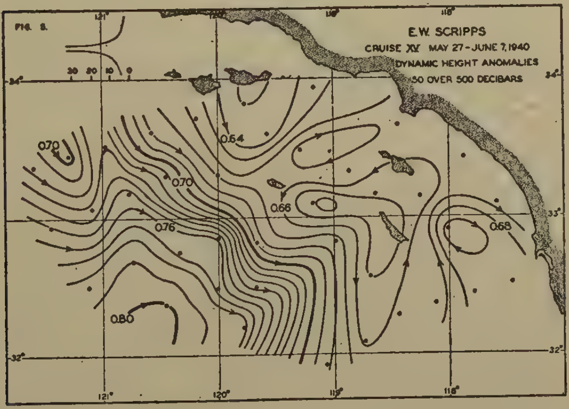


FIG. 5
E.W. SCRIPPS
CRUISE XV MAY 27-JUNE 7, 1940
DYNAMIC HEIGHT ANOMALIES
50 OVER 500 DECIBARS
121°
120°
119°
118°
34°
33°
32°
30
20
10
0.70
0.64
0.66
0.68
0.76
0.80

Figures 4, 5. Dynamic Height Anomalies (current boundaries) 50 over 500 Decibars, 1940.

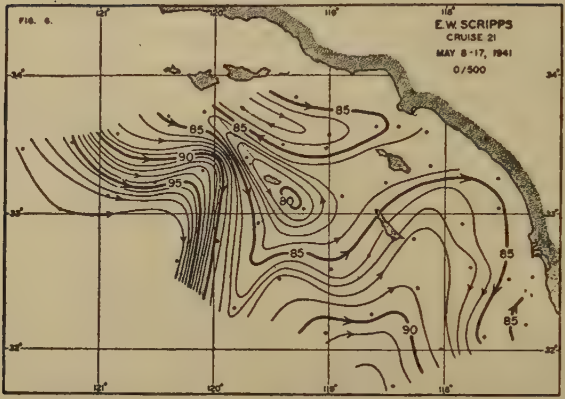


FIG. 6.
E.W. SCRIPPS
CRUISE 21
MAY 8-17, 1941
0/500
121°
120°
119°
118°
34°
33°
32°
85
90
95
80

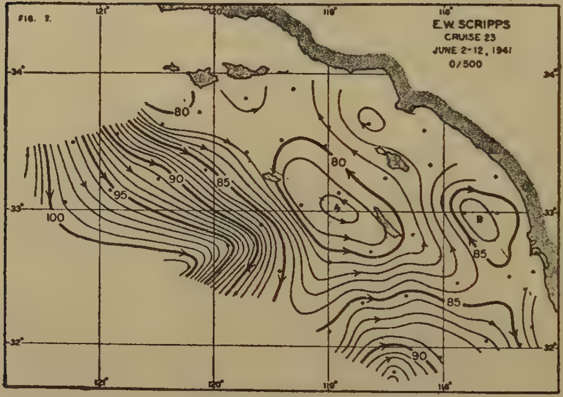


FIG. 7.
E.W. SCRIPPS
CRUISE 23
JUNE 2-12, 1941
0/500
121°
120°
119°
118°
34°
33°
32°
80
85
90
95
100
A
B

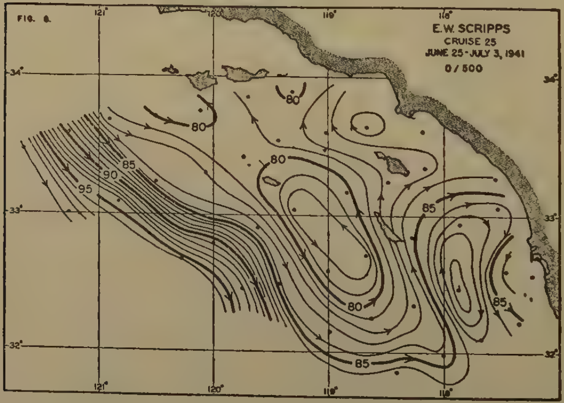


FIG. 8.
E.W. SCRIPPS
CRUISE 25
JUNE 25-JULY 3, 1941
0/500
121°
120°
119°
118°
34°
33°
32°
80
85
90
95

Figures 6–8. Dynamic Height Anomalies (current boundaries) 0 over 500 Decibars, 1941.

level has the same general characteristics, but the position of maximum velocity shifts from 10 to nearly 30 kilometers at the last stage.

THE CONSERVATION OF ANGULAR MOMENTUM, AND RATE OF ENERGY DISSIPATION IN THE EQUIVALENT CIRCULAR EDDIES

The empirical curves drawn to show the relation of velocity to distance from the center of an eddy at its successive stages are only

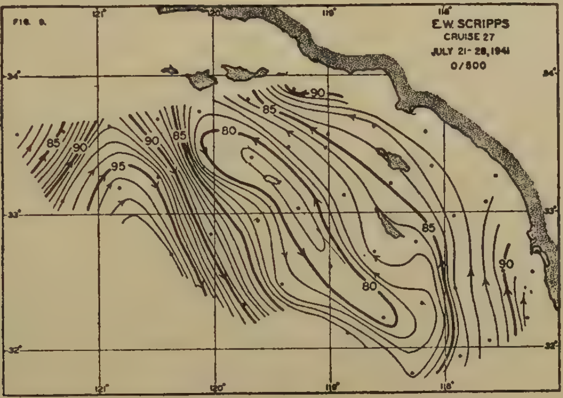


FIG. 9.
121°
120°
119°
118°
E.W. SCRIPPS
CRUISE 27
JULY 21-28,1941
0/500
34°
33°
32°
90
85
80
95
90
85
80

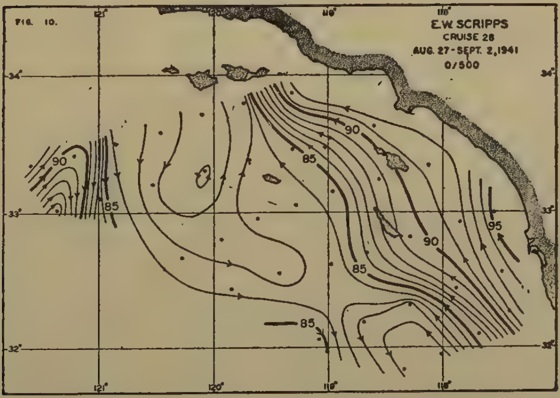


FIG. 10.
121°
120°
119°
118°
E.W. SCRIPPS
CRUISE 28
AUG. 27-SEPT. 2,1941
0/500
34°
33°
32°
90
85
90
85
95
90
85
85

Figures 9, 10. Dynamic Height Anomalies (current boundaries) 0 over 500 Decibars, 1941.

fair approximations to what might result from more detailed observations. Accordingly, only crude estimates of the quantities required can be obtained from these curves. Yet these curves must serve as a

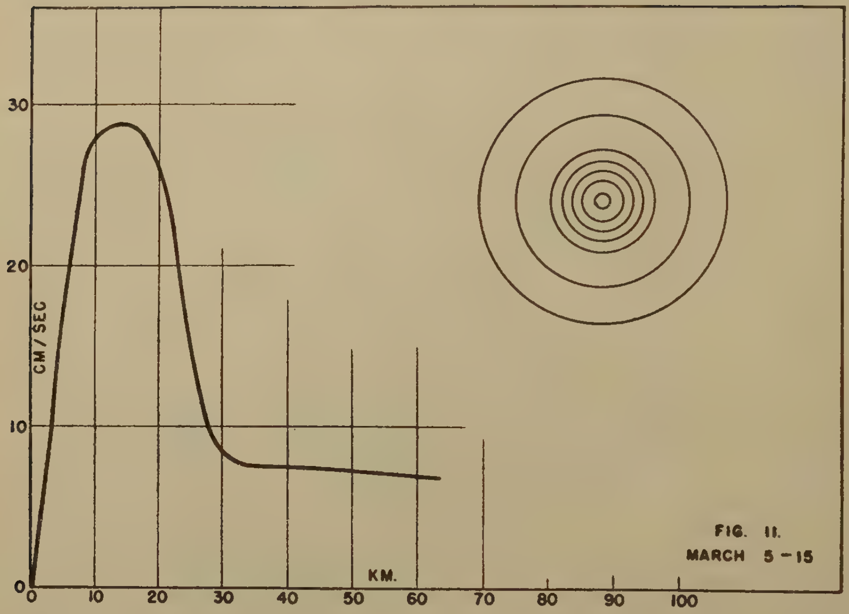


FIG. 11.
MARCH 5 – 15

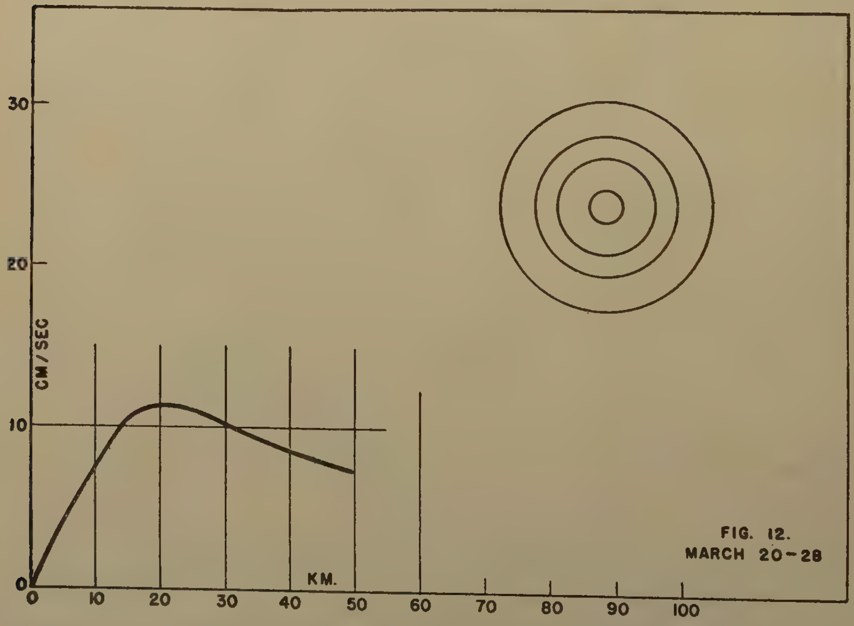


FIG. 12.
MARCH 20 – 28

Figures 11, 12. Circular models of eddies indicated on charts of Dynamic Height Anomalies, 0 over 500 Decibars, and relation of velocity to distance from center, 1940.

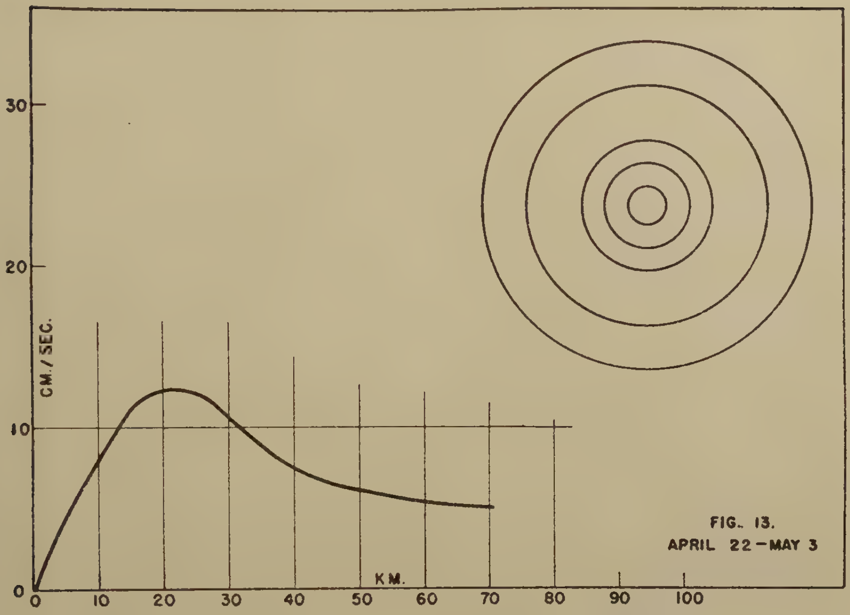


FIG. 13.
APRIL 22—MAY 3

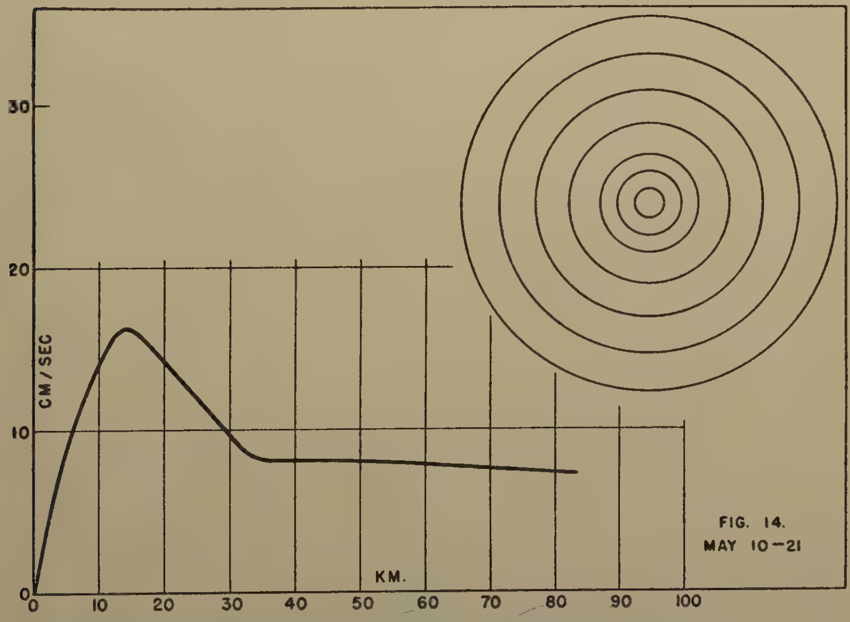


FIG. 14.
MAY 10—21

Figures 13, 14. Circular models of eddies indicated on charts of Dynamic Height Anomalies, 0 over 500 Decibars, and relation of velocity to distance from center, 1940.

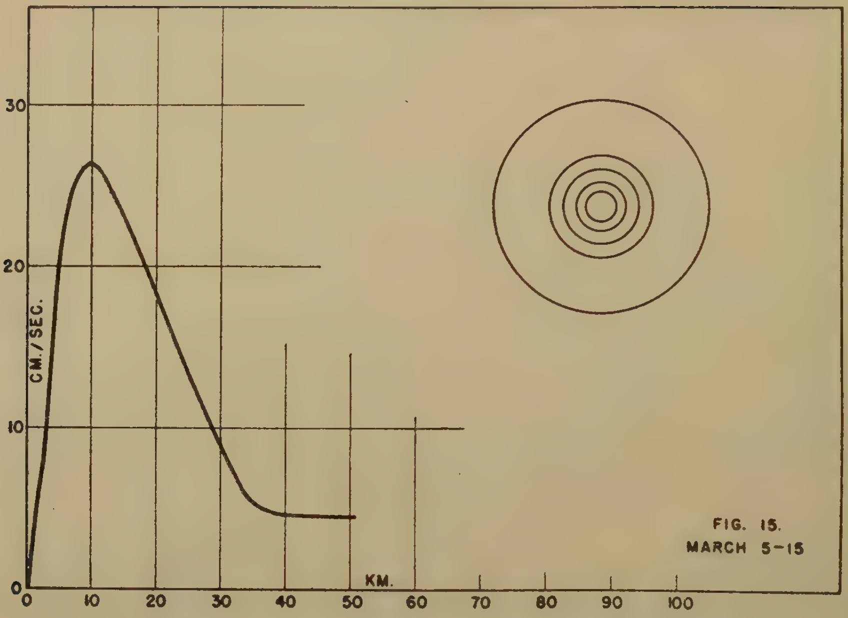


FIG. 15.
MARCH 5–15

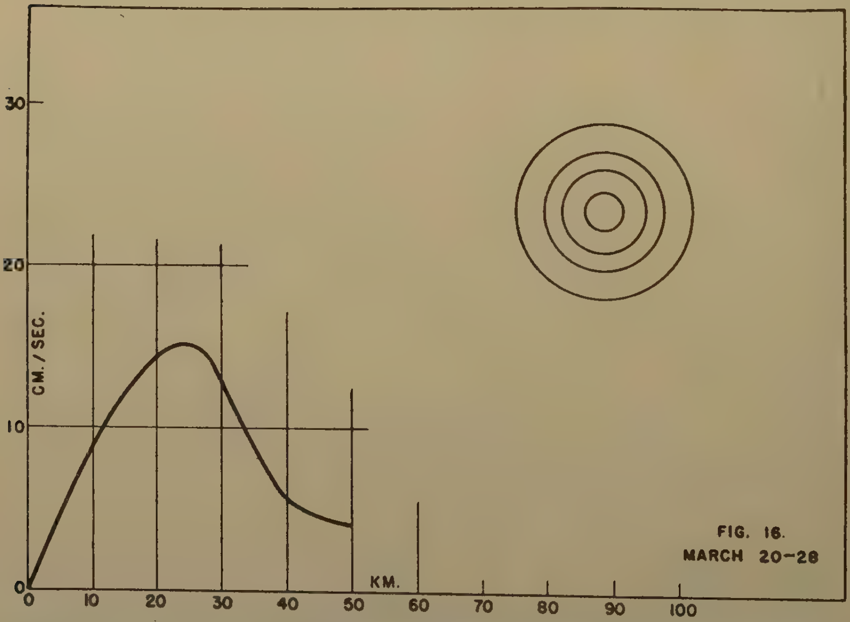


FIG. 16.
MARCH 20–28

Figures 15, 16. Circular models of eddies indicated on charts of Dynamic Height Anomalies, 50 over 500 Decibars, and relation of velocity to distance from center, 1940.

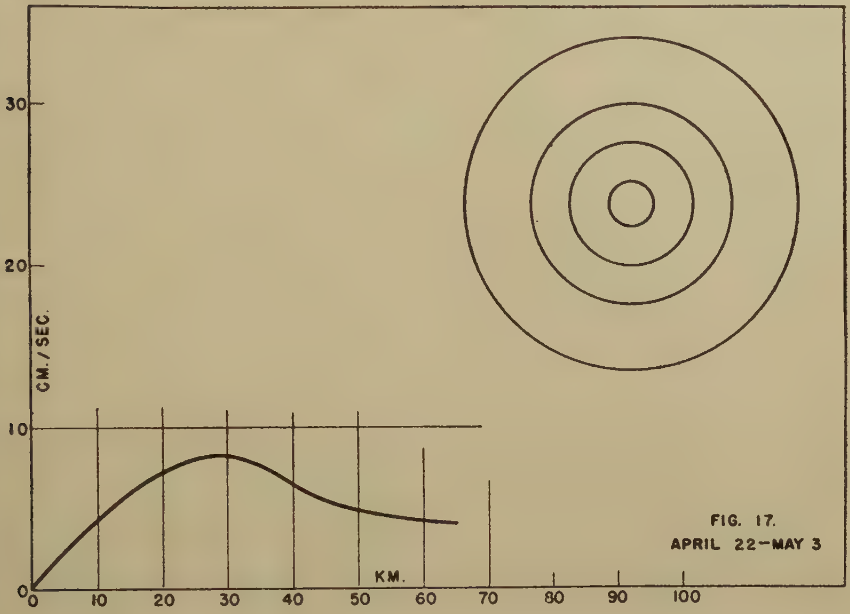


FIG. 17.
APRIL 22—MAY 3

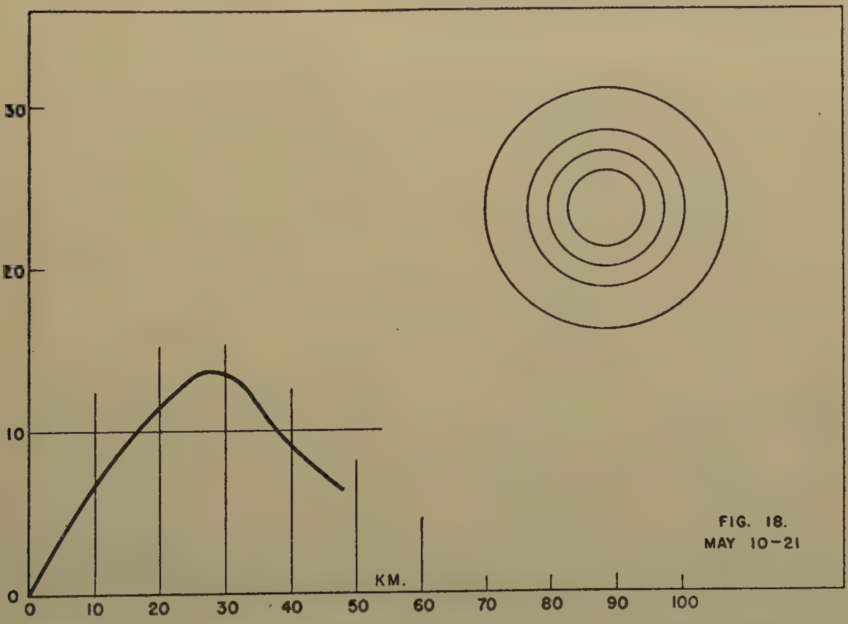


FIG. 18.
MAY 10—21

Figures 17, 18. Circular models of eddies indicated on charts of Dynamic Height Anomalies, 50 over 500 Decibars, and relation of velocity to distance from center, 1940.

basis for the first approach to the dynamics of the eddies, and by making several estimates of each quantity the errors may be reduced.

Since both the velocity and the velocity gradient decrease as the distance, r, from the center increases beyond that corresponding to the maximum velocity, external frictional forces may be neglected, if a sufficiently large portion of the eddy is used as a basis for computations.

Assuming only internal forces, the principle of the conservation of angular momentum applies to the eddy, considered as a mechanical system. Therefore, assuming unit density of the water, the angular momentum of unit thickness of the eddy,

$$\text{A. M.} = 2\pi \int_0^{r'} r^2 q dr \,, \tag{1}$$

should be constant at successive times, where q equals the velocity at distance r. A reasonable length of the maximum radius, r', of the eddies was determined by inspection of the charts and graphs to be 50 kilometers. If the same length is used for the successive stages of each series, a rough approximation will serve the purpose. Values of the angular momentum in c. g. s. units, computed by numerical integration of r^2q read from the empirical curves, are compiled in Table I.

TABLE I. ANGULAR MOMENTUM, IN C. G. S. UNITS, OF HORIZONTAL EDDIES AT THE SURFACE AND AT 50-METER LEVELS

1940	*0/500*	*50/500*	*Sum*
A	3.16×10^{21}	1.81×10^{21}	4.97×10^{21}
14 days			
B	2.46×10^{21}	2.14×10^{21}	4.60×10^{21}
34 days			
C	2.29×10^{21}	1.81×10^{21}	4.10×10^{21}
18 days			
D	2.76×10^{21}	2.24×10^{21}	5.00×10^{21}

1941	*0/500*
A	4.85×10^{21}
48 days	
B	4.81×10^{21}
62 days	
C	2.47×10^{21}

Considering the rather crude approximation of the empirical graph of q against r, the tabulated results agree well with the principle of the conservation of angular momentum.

Values of the kinetic energy, in c. g. s. units, computed for the same eddies by means of the formula

$$\text{K. E.} = 2\pi \int_0^{r'} r \left(\frac{q^2}{2} \right) dr \,, \tag{2}$$

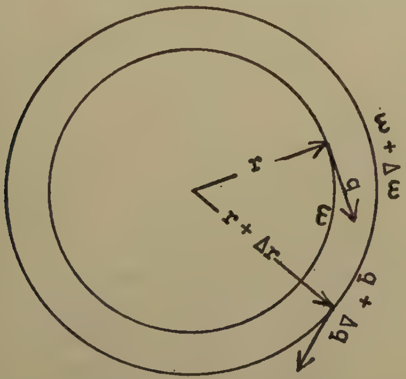
are compiled in Table II.

TABLE II. KINETIC ENERGY, IN C. G. S. UNITS, OF HORIZONTAL EDDIES AT THE SURFACE AND AT 50-METER LEVELS

1940	*0/500*	*50/500*	*Sum*
A	8.7×10^{15}	4.9×10^{15}	13.6×10^{15}
B	3.5×10^{15}	4.5×10^{15}	8.0×10^{15}
C	3.7×10^{15}	1.8×10^{15}	5.5×10^{15}
D	5.5×10^{15}	3.5×10^{15}	9.0×10^{15}

1941	*0/500*
A	7.63×10^{15}
B	6.57×10^{15}
C	1.62×10^{15}

Although the results are somewhat irregular, as was to be expected, there is a general decrease of the kinetic energy in the successive stages of each eddy, and estimates of the average rate of energy dissipation corresponding to various time intervals can easily be computed. The rates of loss of energy can also be computed from formulae involving velocity gradients and the coefficient of eddy viscosity a^2. If the motion is along parallel straight lines, the rate of energy dissipation per unit volume due to friction (Dryden, Murnaghan and Bateman, 1932: 157; Kalinske, 1939: 1394; 1940: 1236) can easily be shown to have the value $a^2\left(\frac{du}{dy}\right)^2$, where u equals the velocity in a direction perpendicular to y. The corresponding expression for motion in concentric circles, assuming circular symmetry of the velocity field, may be derived as follows: Referring to the diagram, where ω equals the angular velocity in radians per second,



$$(\omega + \Delta\omega)\ (r + \Delta r) = \omega r + \omega\Delta r + r\Delta\omega + [(\Delta\omega)\ (\Delta r)] = q_{r+\Delta r}\ ,$$

$$\omega r = q_r\ ,$$

and $$q_{r+\Delta r} - q_r = \omega\Delta r + r\Delta\omega + [(\Delta\omega)\ (\Delta r)]\ .$$

Therefore,

$$\frac{dq}{dr} = \omega + r\frac{d\omega}{dr} \text{ and } r\frac{d\omega}{dr} = \frac{dq}{dr} - \omega = \frac{dq}{dr} - \frac{q}{r}\ . \tag{3}$$

Since $r\Delta\omega$ equals that part of the velocity in the larger circle relative to that in the smaller circle, due to the change in angular velocity with reference to r, the frictional stress due to the relative velocity,

$$\left(r\frac{d\omega}{dr}\right)\Delta r = r\Delta\omega\ ,$$

is

$$\frac{a^2\left(r\dfrac{d\omega}{dr}\right)\Delta r}{\Delta r}\ .$$

Therefore, the power loss per unit horizontal surface area is

$$\left[\frac{a^2\left(r\dfrac{d\omega}{dr}\right)\Delta r}{\Delta r}\right]\left[r\frac{d\omega}{dr}\Delta r\right]\ .$$

Dividing this product by Δr gives the rate of dissipation of energy per unit volume, which reduces to

$$a^2\left(r\frac{d\omega}{dr}\right)^2 = a^2\left(\frac{dq}{dr} - \frac{q}{r}\right)^2\ .$$

Therefore, the rate of energy dissipation of an eddy of unit thickness is

$$2\pi a^2\int_0^{r'} r\left(\frac{dq}{dr} - \frac{q}{r}\right)^2 dr\ , \tag{4}$$

which can be found by numerical integration of the integrand, whose value can be estimated from the empirical graphs of q against r. Finally, by equating the two expressions for the average rates of energy loss,

$$\left.2\pi a^2\int_0^{r'} r\left(\frac{dq}{dr} - \frac{q}{r}\right)^2 dr\ \right]_{t_0,\ t_1}$$

$$= 2\pi \frac{\left.\int_0^{r'} r\left(\frac{q^2}{2}\right) dr\right]_{t_1} - \left.\int_0^{r'} r\left(\frac{q^2}{2}\right) dr\right]_{t_0}}{(t_1 - t_0)}, \quad (5)$$

we obtain a formula for computing the coefficient of eddy viscosity,

$$a^2 = \frac{\left\{\left.\int_0^{r'} r\left(\frac{q^2}{2}\right) dr\right]_{t_1} - \left.\int_0^{r'} r\left(\frac{q^2}{2}\right) dr\right]_{t_0}\right\}}{(t_1 - t_0) \left.\int_0^{r'} r\left(\frac{dq}{dr} - \frac{q}{r}\right)^2 dr\right]_{t_0 t_1}}, \quad (6)$$

where the integral in the denominator should be averaged over the interval t_1 to t_0. A sufficiently close approximation for the present purpose is the arithmetic mean of the values of the integral corresponding to the times t_0 and t_1. The values of a^2 thus found from various combinations of kinetic energy values in Table II are presented in Table III.

TABLE III. VARIOUS ESTIMATES OF THE COEFFICIENT OF EDDY VISCOSITY a^2 IN C. G. S. UNITS

1940	*0/500*	*50/500*
A, B	4.9×10^5	0.3×10^5
A, C	1.8×10^5	1.3×10^5
A, D	0.8×10^5	0.4×10^5
A, E	1.9×10^5	1.3×10^5
1941	*0/500*	
A, B	0.7×10^5	
A, C	2.5×10^5	
B, C	5.7×10^5	

The range is rather large for each of the three series due to the crude empirical relation between q and r. Adopted median values of a^2 for later use are listed in Table IV.

TABLE IV. SELECTED MEDIAN VALUES OF THE COEFFICIENT OF EDDY VISCOSITY FOR EACH OF THREE SERIES OF EDDIES

Year	*Depths*	a^2
1940	0/500	1.5×10^5
1940	50/500	1.0×10^5
1941	0/500	3.0×10^5

PRELIMINARY STUDY OF LARGE EDDIES BASED ON VORTEX-FILAMENT THEORY

The foregoing evidence presented in support of the thesis appears to justify an attempt to predict, from the observed velocity distribution at the earliest date in each series, the relation of velocity to the distance from the center at the successive times of later surveys. This hydrodynamical problem will now be considered in the light of the resemblance of the observed velocity distribution in an eddy to the velocity distribution around a vortex-filament. For a detailed treatment of the vortex-filament, see Lamb (1945: 202–249), Prandtl (1935: 68), and Dryden, Murnaghan and Batemen (1932: 212–218). Vortex lines have the direction of the instantaneous axis of rotation of the fluid. A vortex-tube is formed by the collection of vortex lines perpendicular to a small closed curve, and the fluid within such a tube constitutes a vortex-filament. The circulation, Γ, of a closed curve is defined as the line integral of the tangential component of the velocity around the curve. Twice the angular velocity of the fluid elements is called the vorticity, and the product of the vorticity by the cross section of a vortex-filament equals the circulation which is constant for all cross sections of the tube. In this paper the vorticity will be denoted by ζ, corresponding to the vertical direction of the axis.

There is an instructive analogy between the dissolution of a vortex-filament generated in a viscous fluid and the conduction of heat from a small heated portion of a flat plate. The constant circulation of the vortex corresponds to a constant quantity of heat initially concentrated at a point. This quantity of heat will spread outward from the point by conduction in accordance with the following function,

$$S = \frac{K}{t} \exp\left(-r^2/4\nu t\right), \tag{7}$$

which is a well known solution of the following differential equation of heat conduction, assuming circular symmetry about the point of concentration:

$$\frac{\partial S}{\partial t} = \nu\left(\frac{\partial^2 S}{\partial r^2} + \frac{1}{r}\frac{\partial S}{\partial r}\right); \tag{8}$$

S = amount of heat per unit area;
ν = coefficient of heat conduction;
r = distance from center;
t = time;
K = a constant.

The total amount of heat, within a circle of radius r, equals

$$2\pi \int_0^r rSdr = 4\nu\pi K\,[1 - \exp(-r^2/4\nu t)]\,, \tag{9}$$

which approaches the constant $4\nu\pi K$ as $r \to \infty$, for all values of t. The same limit is approached as $t \to 0$, for all values of r.

Leaving the problem of heat conduction, consider the analogous one of vorticity,

$$\Gamma = 2\pi \int_0^r K/t \exp(-r^2/4a^2t)\,rdr = 2\pi \int_0^r \zeta r dr$$

$$= 4a^2\pi K\,[1 - \exp(-r^2/4a^2t)], \tag{10}$$

where a^2 equals the coefficient of viscosity.

As $r \to \infty$, $\Gamma \to$ the constant $4a^2\pi K = \Gamma_1$, the circulation of an infinitely large circle. Therefore,

$$\Gamma = \Gamma_1\,[1 - \exp(-r^2/4a^2t)]\,. \tag{11}$$

The velocity, q, equals the circulation divided by the circumference:

$$q = \frac{\Gamma_1}{2\pi r}\,[1 - \exp(-r^2/4a^2t)]\,. \tag{12}$$

As $(r^2/4a^2t) \to \infty$, $q \to q_1 = \dfrac{\Gamma_1}{2\pi r}$ equals the usual circulatory motion with a potential. For small radial distances,

$$q = \frac{\Gamma_1 r}{2\pi}\,\{\tfrac{1}{4}a^2t - \tfrac{1}{2}(r/4a^2t)^2 + \ldots\}\,. \tag{13}$$

Since $\zeta = (K/t)\exp(-r^2/4a^2t)$ from the heat conduction analogy, the vorticity equals

$$\zeta = \frac{\Gamma_1}{4a^2\pi t}\exp(-r^2/4a^2t)\,. \tag{14}$$

Therefore, on the axis

$$\zeta = \zeta_1 = \frac{\Gamma_1}{4a^2\pi t}\,, \tag{15}$$

and for very small radii, the velocity equals

$$q = \frac{\Gamma_1 r}{2\pi\,(4a^2t)} = \frac{\zeta_1}{2}r = \frac{\zeta_1 r_1}{2}\theta, \text{ where } r = r_1\theta\,. \tag{16}$$

In general,

$$q = \frac{\Gamma_1}{2\pi r}\left[1 - \exp\left(-r^2/4a^2t\right)\right] = \frac{\zeta_1 r_1}{2}\left[\frac{1 - \exp\left(-\dfrac{\theta^2 r_1^2}{4a^2t}\right)}{\theta}\right]. \quad (17)$$

Accordingly for $t > 0$, the velocity around a vortex-filament is proportional to the distance from the center for small values of r. As r increases, the velocity is inversely proportional to r. Thus the velocity increases from zero to a maximum at some small value of r, and then decreases, approaching zero, as r increases indefinitely. However, as $t \rightarrow 0$, $\zeta_1 \rightarrow \infty$, and the position of maximum velocity approaches the axis. Thus, there is a general qualitative agreement between the velocity distribution around the vortex-filament and that observed in the eddies in the ocean. However, the initial position of the maximum velocity on the axis and the concentration of a very great vorticity in a very small area together constitute insuperable objections to the use of a vortex-filament as a model for detailed quantitative application to a large eddy in the sea. Yet, the general results already found suggest that it would be worth while to attempt the application of the fundamental concept of the vortex-filament to the eddy problem in such a way as to avoid the foregoing objections.

VORTICITY DISTRIBUTION OVER A CIRCULAR AREA AS A SATISFACTORY MODEL OF A LARGE EDDY IN THE OCEAN

In order to take advantage of the fundamental physical properties of the vortex-filament, and to avoid the objections referred to in the foregoing section, assume an initial distribution of vorticity over a circular area of appropriate radius, r_1. Moreover, in this attempt to devise a model of an eddy in the ocean, it will be necessary to place the whole theory on a sound physical basis, defined by the following accepted general hydrodynamical equations of motion in the sea (Sverdrup, Johnson and Fleming, 1942: 433; Sverdrup, 1942: 93, 96). In a left-hand rectangular coordinate system with positive z-axis the general hydrodynamical equations of motion in the ocean for the northern hemisphere are:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = 2\Omega\left(\sin\phi\right)v - \alpha\frac{\partial p}{\partial x} + \alpha R_x\,, \quad (19)$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} = -\,2\Omega\left(\sin\phi\right)u - \alpha\frac{\partial p}{\partial y} + \alpha R_y\,, \quad (20)$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} = -\,2\Omega\,(\sin\phi)\,v_E + g - \alpha\frac{\partial p}{\partial z} + \alpha R_z\,. \quad (21)$$

$u, v, w,$	the velocity components along the coordinate axes;
v_E	the horizontal velocity toward the east;
$\Omega,$	angular velocity of earth's rotation $= .729 \times 10^{-3}$;
$\phi,$	geographic latitude;
$\alpha,$	specific volume of the water;
$p,$	pressure;
$R_x, R_y, R_z,$	components of the frictional force per unit volume;
$g,$	acceleration of gravity.

For the southern hemisphere, change the signs of the first terms on the right-hand side of the first two equations.

Consider only horizontal motion and assume small variations of latitude. Let $2\,\Omega \sin\phi$ = a constant average value, λ, and replace the symbols R, denoting fluid friction, by $a^2\nabla^2 u$, and $a^2\nabla^2 v$, thus obtaining the Navier-Stokes equations of motion in the sea (Sverdrup, Johnson and Fleming, 1942: 470):

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -\,\alpha\frac{\partial p}{\partial x} + \lambda v + a^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), \quad (22)$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -\,\alpha\frac{\partial p}{\partial y} - \lambda u + a^2\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right). \quad (23)$$

(a^2 = kinematic coefficient of eddy viscosity, assumed to be constant.) Take the negative partial derivative of equation (22) with respect to y and the positive partial derivative of equation (23) with respect to x, and add the results:

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + u\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + v\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + \frac{\partial u}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$
$$+ \frac{\partial v}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) = a^2\left[\frac{\partial}{\partial x^2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial y^2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\right]. \quad (24)$$

Substitute ζ for the vorticity, $\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$ = curl of the velocity (Lamb, 1945: 202), and assume the equation of continuity,

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0\,,$$

thus obtaining

$$\frac{\partial \zeta}{\partial t} = a^2\left(\frac{\partial^2 \zeta}{\partial x^2} + \frac{\partial^2 \zeta}{\partial y^2}\right) - \left(u\frac{\partial \zeta}{\partial x} + v\frac{\partial \zeta}{\partial y}\right). \quad (25)$$

Equations for transforming from rectangular to polar coordinates, assuming circular symmetry or $\partial/\partial\theta$ and $\partial^2/\partial\theta^2 = 0$ (Gibson, 1931: 113–115; Woods, 1926: 72–73), are

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta\,, \quad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\theta\,, \quad \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial r^2}\cos^2\theta + \frac{\partial f}{\partial r}\frac{\sin^2\theta}{\text{r}}\,,$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial r^2}\sin^2\theta + \frac{\partial f}{\partial r}\frac{\cos^2\theta}{r}\,, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\sin 2\theta}{2}\left(\frac{\partial^2 f}{\partial r^2} - \frac{1}{r}\frac{\partial f}{\partial r}\right).$$

For circular symmetry,

$$\left(u\frac{\partial\zeta}{\partial x} + v\frac{\partial\zeta}{\partial y}\right)$$

reduces to

$$u\frac{\partial\zeta}{\partial r}\cos\theta + v\frac{\partial\zeta}{\partial r}\sin\theta = -q\frac{\partial\zeta}{\partial r}(\sin\theta\cos\theta - \cos\theta\sin\theta) = 0\,,$$

since $u = -q\sin\theta$ and $v = q\cos\theta$.
Therefore, the final equation is

$$\frac{\partial\zeta}{\partial t} = a^2\left(\frac{\partial^2\zeta}{\partial x^2} + \frac{\partial^2\zeta}{\partial y^2}\right). \qquad (26)$$

For circular symmetry

$$\frac{\partial^2\zeta}{\partial x^2} = \frac{\partial^2\zeta}{\partial r^2}\left(\frac{x^2}{x^2+y^2}\right) + \frac{\partial\zeta}{\partial r}\frac{y^2}{(x^2+y^2)\sqrt{x^2+y^2}}\,,$$

$$\frac{\partial^2\zeta}{\partial y^2} = \frac{\partial^2\zeta}{\partial r^2}\left(\frac{y^2}{x^2+y^2}\right) + \frac{\partial\zeta}{\partial r}\frac{x^2}{(x^2+y^2)\sqrt{x^2+y^2}}\,.$$

Substituting these expressions and simplifying gives the equation of diffusion of vorticity in the sea,

$$\frac{\partial\zeta}{\partial t} = a^2\left(\frac{\partial^2\zeta}{\partial r^2} + \frac{1}{r}\frac{\partial\zeta}{\partial r}\right), \qquad (27)$$

which is the same form as the equation of heat conduction (8). Thus the analogy of vorticity diffusion to heat conduction used in the section on the vortex-filament is established for motion in the sea.

If the coefficient of eddy viscosity is not restricted to the constant value a^2, but is assumed to be μ, a function of x and y, a method similar to the foregoing yields the equation

$$\frac{\partial\zeta}{\partial t} = \mu\left(\frac{\partial^2\zeta}{\partial r^2} + \frac{1}{r}\frac{\partial\zeta}{\partial r}\right) + \zeta\left(\frac{\partial^2\mu}{\partial r^2} + \frac{1}{r}\frac{\partial\mu}{\partial r}\right) + 2\frac{\partial\mu}{\partial r}\frac{\partial\zeta}{\partial r}\,, \qquad (28)$$

where μ is a function of r.

Having derived equation (27) of the diffusion of vorticity in the sea, consider initial conditions adapted to a large eddy:

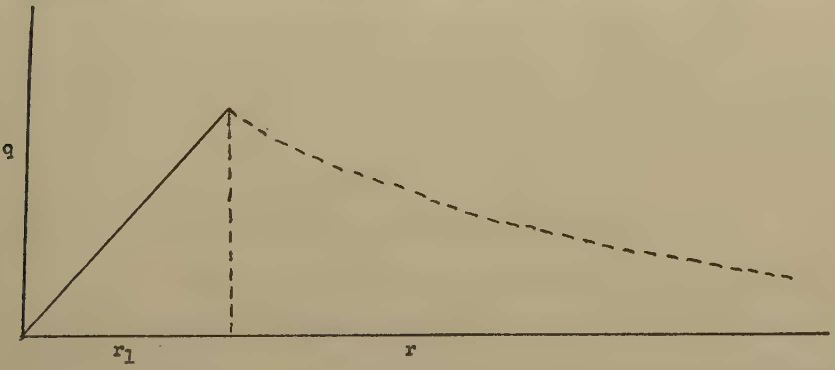
$$\zeta = \zeta_1\,, \quad r \leqq r_1\,, \quad t = 0\,,$$
$$\zeta = 0\,, \quad r > r_1\,, \quad t = 0\,,$$

where ζ_1 is constant and r_1 is a constant radius to be determined. Then the initial velocity distribution is given by equations

$$q = \frac{\zeta_1 \int\limits_0^r 2\pi r dr}{2\pi r} = \frac{\zeta_1 r}{2} = \left(\frac{\zeta_1 r_1}{2}\right)\theta\,, \quad \theta = r/r_1 \leqq 1\,, \qquad (29)$$

$$q = \frac{\zeta_1 \int\limits_0^\infty 2\pi r dr}{2\pi r} = \frac{\zeta_1 \int\limits_0^{r_1} 2\pi r dr}{2\pi r} = \frac{\zeta_1 r_1^2}{2r} = \left(\frac{\zeta_1 r_1}{2}\right)\frac{1}{\theta}\,, \quad \theta > 1\,. \qquad (30)$$

In such a model the initial velocity would be proportional to r for $r \leqq r_1$ and inversely to r for $r > r_1$, with a maximum value $\frac{\zeta_1 r_1}{2}$ at the distance r_1, as shown in the following diagram:



The graphs (Figs. 11 and 16), corresponding to the earliest dates of the 1940 surveys, may be regarded as in good agreement with the slight alteration of this diagram, expected to represent the effect of diffusion during the time from the initial conditions to that of the first cruise of the series.

We shall now attempt to find a solution of equation (27) that satisfies the foregoing initial conditions. A well known elementary solution is

$$F(\lambda)\exp\left(-\lambda^2 a^2 t\right) J_0(\lambda r)\,,$$

where $F(\lambda)$ is an arbitrary function of λ and J_0 is a Bessel function of zero order. Summing these elementary solutions with respect to λ, which is independent of t and r, gives the formal solution

$$\zeta = \int_0^\infty F(\lambda)\, J_0\,(\lambda r) \exp\,(-\,\lambda^2 a^2 t)\, d\lambda\,. \tag{31}$$

For $t = 0$, we require that

$$\int_0^\infty F(\lambda)\, J_0\,(\lambda r)\, d\lambda = \begin{cases} \zeta_1\,, r \leqq r_1 \\ 0\;\;, r > r_1 \end{cases}. \tag{32}$$

According to Gray, Mathews and MacRobert (1922: 64–78)

$$\int_0^\infty J_1\,(x)\, J_0\,(\alpha x)\, dx = \begin{cases} 1\,, \alpha^2 < 1 \\ 1/2\,, \alpha^2 = 1 \\ 0\,, \alpha^2 > 1 \end{cases}. \tag{33}$$

Therefore, we readily obtain the required formal solution

$$\zeta = \zeta_1\, r_1 \int_0^\infty J_1\,(\lambda r_1)\, J_0\,(\lambda r) \exp\,(-\,\lambda^2 a^2 t)\, d\lambda\,. \tag{34}$$

This is mathematically the same solution as that investigated by McEwen (1946: 618) in a problem of diffusion of heat or particulate matter from a circular area of concentration. Accordingly, the same transformations of the formal equation to expressions adapted to numerical work will be restated here:

$$\zeta = \zeta_1\,[1 - \exp\,(-\,r_1^2/4a^2t)]\,, \text{ for } r = 0\,, \tag{35}$$

$$\zeta = (\zeta_1/2)\;[1 - I_0\,(r_1^2/2a^2t) \exp\,(-\,r_1^2/2a^2t)]\,, \text{ for } r = r_1, \tag{36}$$

$$\zeta = \zeta_1\,\{[1 - \exp\,(-\,r_1^2/4a^2t)] - (r_1/2a^2t) \exp\,(-\,r_1^2/4a^2t) \int_0^r I_1\,(rr_1/2a^2t) \exp\,(-\,r^2/4a^2t)\, dr\}\,, \tag{37}$$

where $I_i\,(x)$ is the modified Bessel function of the first kind. Also, the function $e^{-x}I_i\,(x)$, denoted by $f_i\,(x)$, has been calculated for $i = 0$ to 23, $x = 0$ to 20; the results are available in photostat copies on file at the Scripps Institution. Accordingly, equation (36) reduces to

$$\zeta = (\zeta_1/2)\;[1 - f_0\,(r_1^2/2a^2t)]\,, \text{ for } r = r_1\,. \tag{38}$$

Equation (37) can be evaluated by numerical integration. However, by repeated application of integration by parts, McEwen (1946: 618) obtained the expansion ζ into the convergent series

$$\zeta = \zeta_1 \left\{ 1 - e^{-\frac{(r-r_1)^2}{4a^2t}} \left[f_0\left(\frac{r\,r_1}{2a^2t}\right) + \left(\frac{r}{r_1}\right) f_1\left(\frac{r\,r_1}{2a^2t}\right) + \left(\frac{r}{r_1}\right)^2 f_2\left(\frac{r\,r_1}{2a^2t}\right) + \ldots \right] \right\} \begin{matrix}\text{for} \\ r < r_1\end{matrix}, \tag{39}$$

and

$$\zeta = \zeta_1 \, e^{-\frac{(r-r_1)^2}{4a^2t}} \left\{ \frac{r_1}{r} f_1\left(\frac{r\,r_1}{2a^2t}\right) + \left(\frac{r_1}{r}\right)^2 f_2\left(\frac{r\,r_1}{2a^2t}\right) + \left(\frac{r_1}{r}\right)^3 f_3\left(\frac{r\,r_1}{2a^2t}\right) + \ldots \right\} \begin{matrix}\text{for} \\ r > r_1\end{matrix}. \tag{40}$$

Having thus obtained a series of values of ζ as a function of r corresponding to any value of t, the velocity q can be obtained from

$$q = \frac{\int_0^r \zeta r dr}{r} \tag{41}$$

by numerical integration. The corresponding formal equation of velocity is

$$q = \zeta_1 r_1 \int_0^\infty \frac{J_1(\lambda r_1)\, J_1(\lambda r)}{\lambda} \exp(-a^2t\lambda^2)\, d\lambda, \tag{42}$$

which can be reduced to

$$q = \frac{\zeta_1 r_1}{2} \left\{ 1 - [f_0(r_1^2/2a^2t) + f_1(r_1^2/2a^2t)] \right\} \tag{43}$$

for $r = r_1$. This equation, which can easily be evaluated from the manuscript tables of $f_0(x)$ and $f_1(x)$, serves as a useful check on the foregoing methods of computation.

For convenience in carrying out the computations corresponding to a series of values of t, required in applications to the observations, the formulae were transformed as follows, using the equation $r = r_1\theta$. Denoting by $\{\ \}_1$ the coefficient of ζ_1, equation (39), and by $\{\ \}_2$ the coefficient of ζ_1, equation (40),

$$q = \left(\frac{\zeta_1 r_1}{2}\right) \frac{\int_0^\theta 2\{\ \}_1\, \theta d\theta}{\theta}, \quad \theta \gtreqless 1, \tag{44}$$

$$q = \frac{\int_0^1 2 \{ \ \}_1 \, \theta d\theta + \int_1^\theta \{ \ \}_2 \, \theta d\theta}{\theta}, \quad \theta > 1 . \tag{45}$$

For any value of t, $r_1^2/4a^2t$ = a constant c. Accordingly,

$$\{ \ \}_1 = 1 - \exp(-c[1-\theta]^2)\,[f_0(2c\theta) + \theta f_1(2c\theta) + \theta^2 f_2(2c\theta) + \ldots], \tag{46}$$

$$\{ \ \}_2 = 2\exp(-c[1-\theta]^2)\,[(1/\theta) f_1(2c\theta) + (1/\theta)^2 f_2(2c\theta) + (1/\theta)^3 f_3(2c\theta) + \ldots]. \tag{47}$$

The values of $\{ \ \}_1$ and $\{ \ \}_2$ can also be computed by means of the single equation (37), using the same transformation, thus:

$$\{[1 - e^{-c}] - 2c \int_0^\theta e^{-c(1-\theta)^2} f_1(2c\theta)\, d\theta\}, \tag{48}$$

which can be evaluated by numerical integration.

PREDICTION OF VELOCITIES IN CIRCULAR EDDIES BY MEANS OF THE FOREGOING MODEL, COMPARED WITH THOSE OBSERVED

Values of the vorticity, ζ, and the relative velocity, q', as functions of the distance, r, were computed by means of the foregoing formulae, corresponding to each of the three values of the coefficient of eddy-viscosity, a^2, found by the energy method, and for times, t, equal to 0, 1, 2, 3, . . . 15 days. The value t_0 of t, for which the computed relative velocities agreed best with the observed relative velocities at the earliest survey date of each of the three series of cruises, was found by comparing computed and observed graphs. This time, t_0, is the estimated time from the initial condition when $\zeta = \zeta_1$ for $r \lesseqgtr r_1$ and $\zeta = 0$ for $r > r_1$, corresponding to the first cruise of the series. In all three cases, the value 6 was selected for t_0. The value of r_1 was found by inspection of the empirical curves of q plotted against r to be 10 kilometers for 1940 and 30 kilometers for 1941 (page 194). In each case a value was assigned to the constant coefficient $(\zeta_1 r_1/2)$ that showed the best agreement between the computed and observed velocities for $t = t_0$. As already shown on pages 207–209, this coefficient is the maximum velocity and corresponds to $r = r_1$ and $t = 0$. For the surface, and the 50-meter level, 1940, and the surface, 1941, these coefficients were found to have the respective values: 42, 35.3 and 18 cm/sec. It remains to determine the value of t corresponding to each cruise by adding to the first value, 6, the time from the first to the second cruise, from the second to the third, etc., and to evaluate the formulae for each value of t thus determined, in order to predict the velocity distributions. The three series of such theoretical results shown by the curves, Figs. 19–23, 24–28, and 29–33, are to be compared with the histograms representing corresponding average velocities observed between the indicated values of r. The general agreement between theory and observation appears to be sufficiently close to prove the thesis offered at the beginning.

CONCLUDING REMARKS

Consideration of the problem of large eddies as a whole leads to calling attention to certain points not specifically mentioned before.

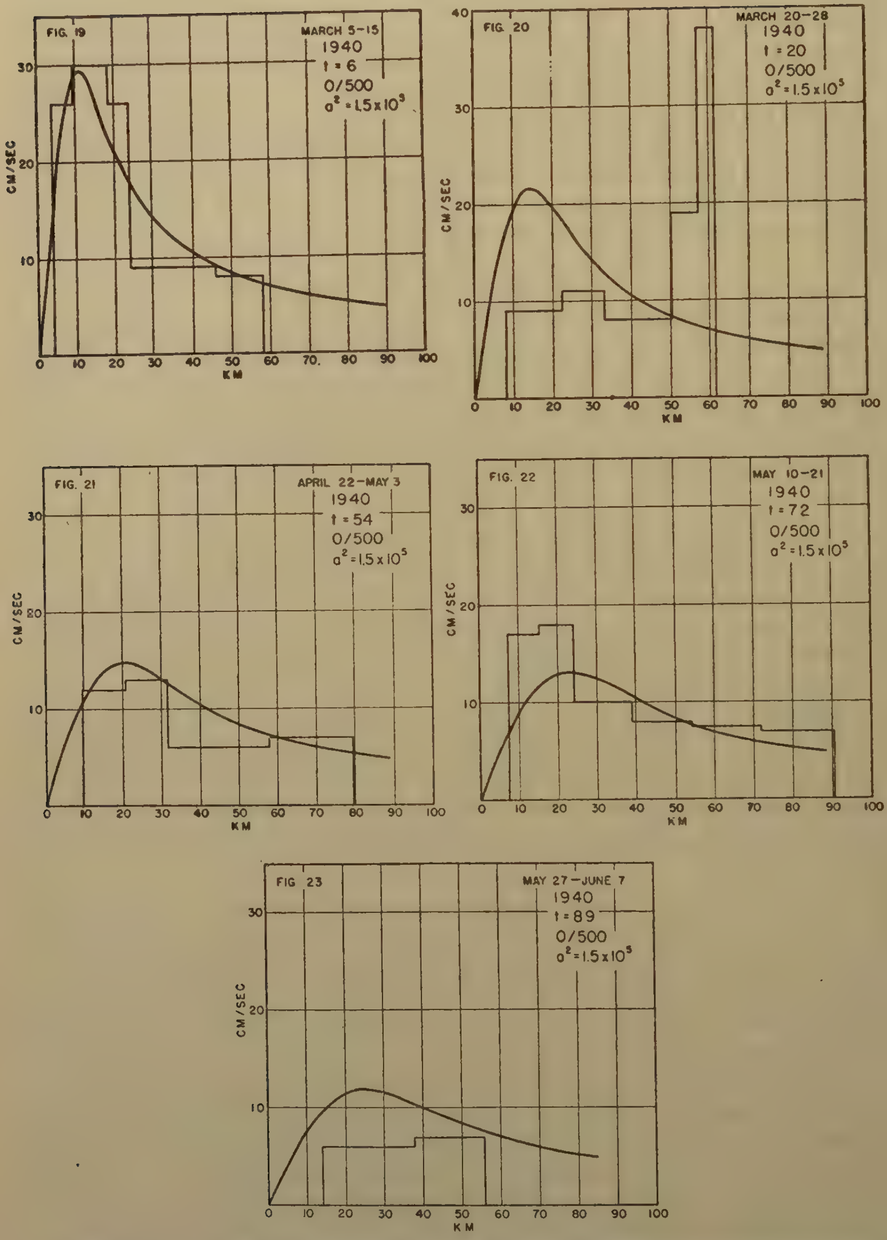
First, the basic hydrodynamic equations (page 204) contained no term representing the wind stress. This omission is justified because the eddies were shown to agree with the principle of the conservation of angular momentum, thus implying that the resultant torque of the wind stress was zero.

Second, the equation of vorticity implies that the shearing stress is due to the horizontal transport of available initial vorticity. This conservative property of vorticity also implies no solenoids in the horizontal plane, no vertical velocity or a uniform distribution of vertical velocity, and no horizontal convergence or divergence (Petterssen and Austin, 1942: 17–19).

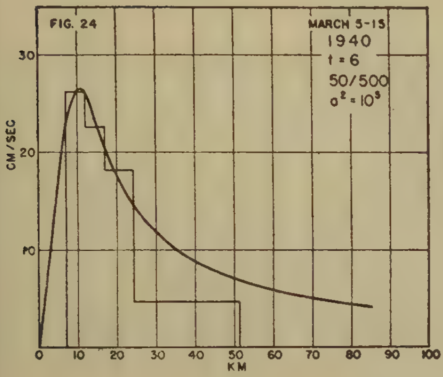
Third, in a circular vortex the fluid elements or particles move in circular orbits around a common axis and all particles lying on one and the same circle have one and the same velocity. In the limiting case of steady motion, the circles around the axis are likewise the paths of the particles and the lines of flow represent the field of velocity (Bjerknes, 1921: 31).

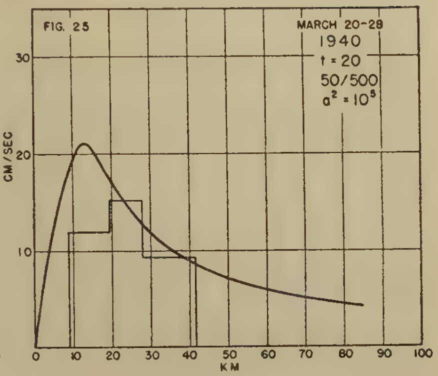
Fourth, it follows from equation (41) that $\frac{q}{r} + \frac{dq}{dr} = \zeta$. From this, the integrand in equation (4) of the rate of dissipation of kinetic energy can be expressed in the form $r\,(\zeta - 2\,q/r)^2$, which is a convenient expression for direct numerical tabulation, in using the theoretical relation of q to r.

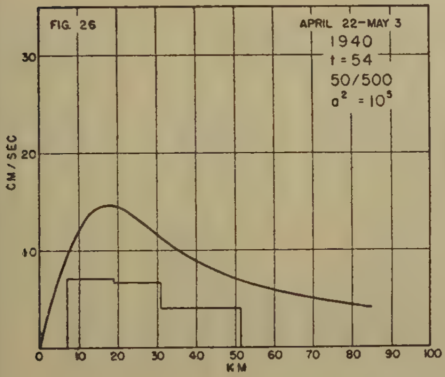
In attempting to apply mathematics to the investigation of any real phenomena, it is necessary to devise a simplified model of the phenomena. In such a model, actual conditions are necessarily simplified and various details are ignored. The success of this kind of an attack depends upon the extent to which essential features of the actual phenomena are preserved, and upon the relative importance of the details ignored. If any elements of the mathematical model depart essen-

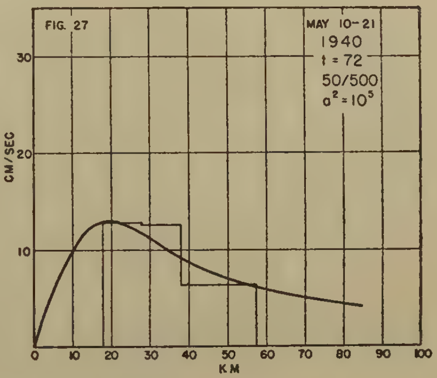


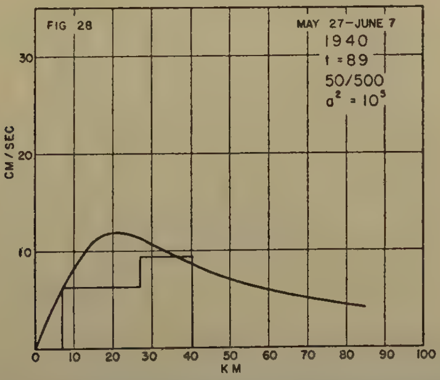
Figures 19–23. Curves of theoretical relation of velocity to distance from center of eddies, and histograms of velocity corresponding to average distance between lines of Dynamic Anomaly curves (current boundary lines). 0 over 500 Decibars. 1940.



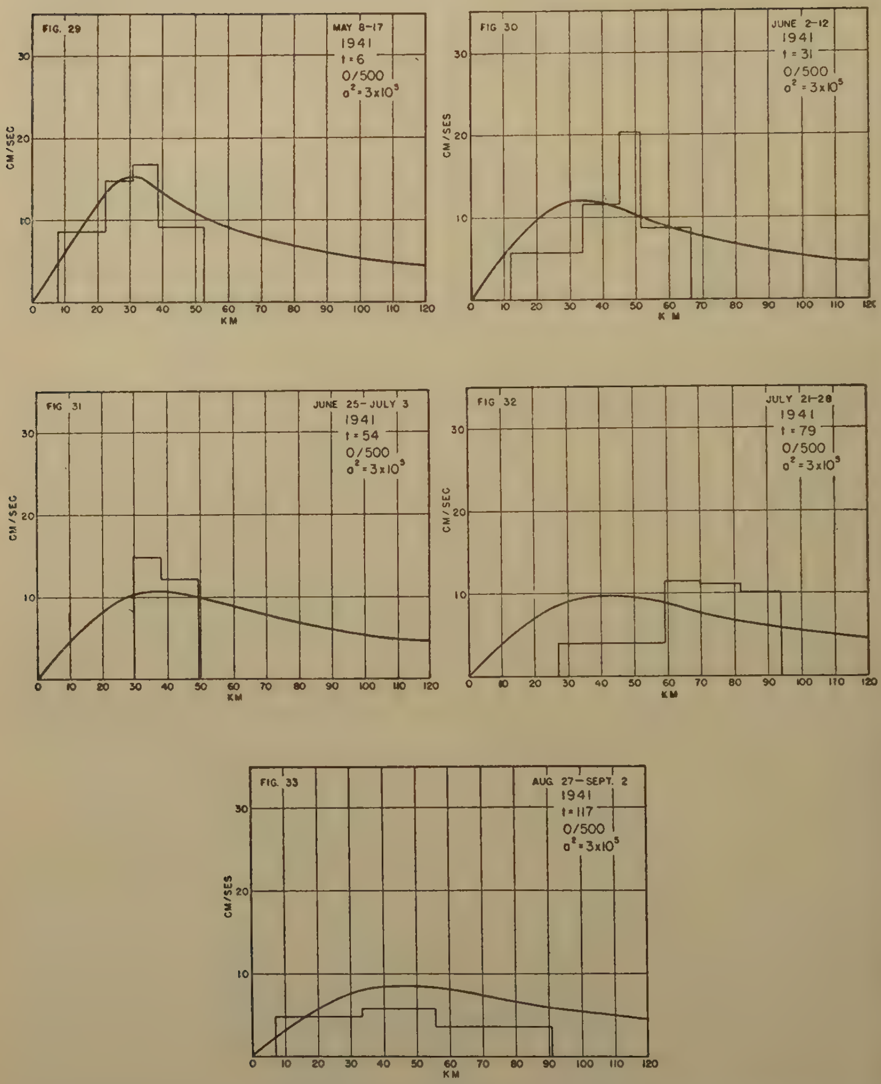








Figures 24–28. Curves of theoretical relation of velocity to distance from center of eddies, and histograms of velocity corresponding to average distance between lines of Dynamic Anomaly curves (current boundary lines), 50 over 500 Decibars, 1940.



Figures 29–33. Curves of theoretical relation of velocity to distance from center of eddies, and histograms of velocity corresponding to average distance between lines of Dynamic Anomaly curves (current boundary lines), 0 over 500 Decibars, 1941.

tially from the reality, the deductions from it may seriously contradict the observations. Since we can not be sure beforehand regarding the adequacy of any model that we may devise, we must test it by making deductions from it and comparing them with observations. Thus, by a trial and error method we may eventually succeed in devising a model sufficiently in accord with reality, yet simple enough to deal with mathematically.

It follows from consideration of the various observed features of the large eddies, and the corresponding deduced results, that, taken as a whole, the evidence is convincing in support of the initialthes is regarding large horizontal eddies in the sea off southern California.

Also, it may be mentioned that, assuming an appropriate initial constant low temperature of a circular area of a 60-kilometer radius, centered at the center of the surface eddy in 1940, the average temperature varied with distance from the center, as would be expected from the law of lateral diffusion. The surface salinity in 1941 also conformed to the same diffusion equation. The coefficients of diffusion obtained in this way were respectively, 1.8×10^7 and $.54 \times 10^7$ c. g. s., which are in good agreement with values for this region found by other methods (Sverdrup, Johnson and Fleming, 1942: 485). These general conclusions imply that the water mass maintains its integrity and position.

REFERENCES

BJERKNES, VILHELM

1921. On the dynamics of the circular vortex with applications to the atmosphere and atmospheric vortex and wave motions. Geofys. Publ., *2* (4): 1–88.

DRYDEN, H. L., F. D. MURNAGHAN AND H. BATEMAN

1932. Report of the Committee on Hydrodynamics, Division of Physical Sciences. Bull. nat. Res. Coun., Wash., No. 84. 634 pp.

GIBSON, G. A.

1931. Advanced Calculus. MacMillan & Co., Ltd., London. xvii + 510 pp.

GRAY, ANDREW, G. B. MATHEWS AND T. M. MACROBERT

1922. A treatise on Bessel functions and their applications to physics. MacMillan & Co., Ltd., London. 2nd ed. xiv + 327 pp.

KALINSKE, A. A.

1939. Relation of the statistical theory of turbulence to hydraulics. Proc. Amer. Soc. civ. Engrs., *65:* 1387–1406.

1940. Relation of the statistical theory of turbulence to hydraulics; Discussion. Proc. Amer. Soc. civ. Engrs., *66:* 1229–1236.

LAMB, HORACE

1945. Hydrodynamics. Dover Publications, New York. sixth ed. xv + 738 pp. (Circular Vortex, viscosity considered. pp. 591–592.)

MCEWEN, G. F.

1946. Diffusion from an instantaneous circular area source (preliminary report). Bull. Amer. math. Soc., *52* (7): 618 (Abstr.).

PETTERSSEN, SVERRE AND J. M. AUSTIN
1942. Fronts and frontogenesis in relation to vorticity. Pap. phys. Oceanogr. Meteorol., *7* (2): 1–37.
PRANDTL, LUDWIG
1935. The mechanics of viscous fluids (pp. 34–208). In: Aerodynamic Theory. W. F. Durand, Editor-in-Chief. *3*. xiv + 354 pp.
SVERDRUP, H. U.
1942. Oceanography for Meteorologists. Prentice-Hall, Inc., New York. xv + 246 pp.
SVERDRUP, H. U., M. W. JOHNSON AND R. H. FLEMING
1942. The Oceans. Prentice-Hall, Inc., New York. x + 1087 pp. (Dimensions and units of terms used in mechanics, p. 401; Equation of continuity, p. 452.)
SVERDRUP, H. U., AND STAFF OF SCRIPPS INSTITUTION OF OCEANOGRAPHY
1938, 1939, 1940, 1941. Oceanographic observations on the "E. W. Scripps" Cruises. University of California Press, Berkeley and Los Angeles.
WOODS, F. S.
1926. Advanced Calculus. Ginn & Company, New York. ix + 397 pp.

THE EFFECT OF OCEAN CURRENTS ON INTERNAL WAVES

By

BERNHARD HAURWITZ

Woods Hole Oceanographic Institution[1]
Woods Hole, Massachusetts
and
College of Engineering, New York University
New York, N. Y.

In a fluid whose density changes with the height, either abruptly at a sharp surface of discontinuity or gradually, it is possible to have internal wave motions whose amplitude is greatest at the density discontinuity or, in the case of a gradual density change, somewhere in the interior of the fluid and not at the free upper surface where the surface waves have their maximum amplitude. The occurrence of such internal waves has been demonstrated by Ekman (1904), Seiwell (1937, 1942), Ufford (1947), and others. A review of our theoretical and empirical knowledge of internal waves has been given by Sverdrup, *et al.* (1942: 585–602). In the theoretical investigations it has always been assumed that the mean velocity of each layer, *i. e.*, the velocity apart from the wave motion, is zero. In many cases, especially where ocean currents are present, this assumption will not be satisfied. It is the purpose of the present paper to investigate whether or not modifications in the internal wave motion would be brought about by the existence of such currents on which the internal wave motion may be superimposed.

The presence of undisturbed currents on which the wave motions are superimposed complicates considerably the expressions for the wave periods, wave velocities, amplitude ratios at different levels, etc. Therefore, a very simple model will be considered which consists of two fluid layers (Fig. 1). The free upper surface (0) may be at the level $z = 0$, the rigid lower surface (2) at the level $z = -H$. Both fluids are separated by an internal boundary at the level $z = -h^{I}$. Each of the two fluids is incompressible and homogeneous, the densities being ρ^{I} and ρ^{II}, respectively, for the upper and lower fluid layers. The velocities in both layers, in the absence of wave motion, are U^{I} and U^{II}, respectively, both parallel to the x-axis. These current

---

[1] Contribution No. 445 from the Woods Hole Oceanographic Institution.

velocities may also be assumed to be constant within each layer and to change abruptly at the surface of discontinuity. In the absence of wave motion the pressure distribution is given by the expression

$$P = g\rho z + \text{const.},$$

where ρ is the density and g the acceleration of gravity. The wave motion may be considered as a small perturbation superimposed on the horizontal current U, and the velocity components in this case may be $U + u$, w while the pressure is $P + p$. Under the assumption that the perturbation quantities are small (designated by small letters), the equations of motion and continuity may be written

$$\begin{aligned}
&\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0\,, \\
(1)\qquad &\frac{\partial w}{\partial t} + U\frac{\partial w}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial z} = 0\,, \\
&\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0\,.
\end{aligned}$$

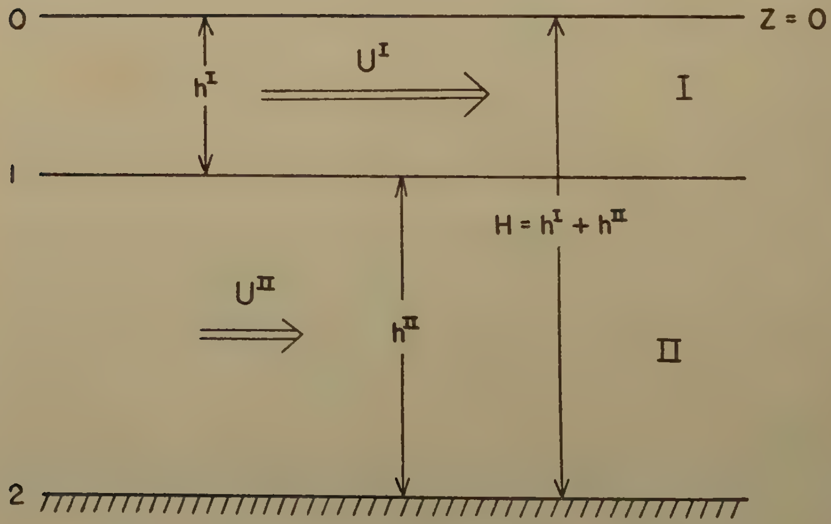


Figure 1. Stratification of fluid.

The equation of continuity is satisfied if a stream function ψ exists, such that

$$u = -\frac{\partial \psi}{\partial z}, \quad w = \frac{\partial \psi}{\partial x}. \tag{2}$$

Since the rigid lower boundary at $z = -H$ must be a stream line, the solution of (1) for the lower layer in the case of waves travelling in the x-direction may be written in the form

$$\psi^{II} = K^{II} \sinh \alpha(z + H)\, e^{i(\alpha x - \beta t)}, \tag{3}$$

$$p^{II} = -\alpha c^{II} \rho^{II} K^{II} \cosh \alpha(z + H)\, e^{i(\alpha x - \beta t)}. \tag{4}$$

Here, K^{II} is an arbitrary constant, $\alpha = 2\pi/L$, where L is the wave length, $\beta = 2\pi/T$, where T is the period, $c = \beta/\alpha$, the wave velocity, and $c^{II} = c - U^{II}$.

For the upper layer, the solution of (1) may be written in the form

$$\psi^{I} = (K_1{}^{I} e^{\alpha z} + K_2{}^{I} e^{-\alpha z})\, e^{i(\alpha x - \beta t)}, \tag{5}$$

$$p^{I} = -\alpha c^{I} \rho^{I} (-K_1{}^{I} e^{\alpha z} + K_2{}^{I} e^{-\alpha z})\, e^{i(\alpha x - \beta t)}, \tag{6}$$

where $K_1{}^{I}$ and $K_2{}^{I}$ are arbitrary constants and where $c^{I} = c - U^{I}$. Since at the free surface the pressure must remain constant,

$$P^{I} + p_{z=0}{}^{I} = -g\rho^{I} z + \alpha c^{I} \rho^{I} (-K_1{}^{I} + K_2{}^{I})\, e^{i(\alpha x - \beta t)} = \text{const.} \tag{7}$$

In the expression for p^{I} the actual height of the free surface has been replaced by the height of the free surface in the undisturbed position, $z = 0$. This simplification involves only an error of higher order. It follows that the equation of the free surface may be written in the form

$$z - Z_0\, e^{i(\alpha x - \beta t)} = \text{const.}, \tag{8}$$

where the amplitude of the free surface

$$Z_0 = \frac{\alpha}{g} c^{I} (-K_1{}^{I} + K_2{}^{I}).$$

The kinematic boundary condition requires that a particle which was once part of the free surface remains in this surface. If in general the equation of the disturbed free surface is

$$F + f = 0,$$

where $F = 0$ denotes the equation for the undisturbed boundary, the kinematic condition becomes

$$\left(\frac{\partial}{\partial t} + U^{I} \frac{\partial}{\partial x}\right) f + w_{z=0}{}^{I} \frac{\partial F}{\partial z} = 0.$$

Here, only an error of higher order is involved if the value of w at the actual disturbed position of the boundary is replaced by the value at the undisturbed position. The last equation leads to the following equation for Z_0,

$$c^I Z_0 = -(K_1{}^I + K_2{}^I)\,.$$

By means of the two expressions for Z_0, (5) and (6) may be written as follows:

$$\text{(9)} \qquad \psi^I = -Z_0\left[\frac{g}{\alpha c^I}\sinh \alpha z + c^I \cosh \alpha z\right] e^{i(\alpha x - \beta t)}$$

$$\text{(10)} \qquad p^I = \alpha c^I \rho^I Z_0\left[\frac{g}{\alpha c^I}\cosh \alpha z + c^I \sinh \alpha z\right] e^{i(\alpha x - \beta t)}\,.$$

The equation of the internal boundary, which, in the absence of wave motion, is $z = -h^I$, is given by the condition that the pressure at this discontinuity must be continuous. There are two kinematic boundary conditions which express the fact that particles in the upper as well as in the lower layer, which at one time were at the boundary, must remain there. Similar to the procedure in the case of the free surface, these conditions permit one to express the arbitrary constant K^{II} in (3) and (4) by the amplitude of the internal boundary Z_1,

$$\text{(11)} \qquad K^{II} = -\frac{Z_1 c^{II}}{\sinh \alpha h^{II}}\,.$$

Furthermore, the ratio of the amplitudes of the free surface Z_0, and of the internal boundary Z_1, is found to be

$$\text{(12)} \qquad \frac{Z_1}{Z_0} = -\frac{g}{\alpha c^I{}_2}\sinh \alpha h^I + \cosh \alpha h^I\,,$$

and the wave velocity c must satisfy the equation

$$\text{(13)} \quad c^{I^4} + \sigma a^I a^{II} c^{I^2} c^{II^2} - \frac{g}{\alpha}\sigma(a^{II}c^{II^2} + a^I c^{I^2}) + \frac{g^2}{\alpha^2}(\sigma - 1) = 0.$$

Here the following abbreviations have been used:

$$\frac{\rho^{II}}{\rho^I} = \sigma, \quad \coth \alpha h^I = a^I, \quad \coth \alpha h^{II} = a^{II}\,.$$

The amplitudes Z_0 and Z_1 of the free and of the internal boundary give the amplitudes of the vertical motion at the levels $z = 0$ and $z = -h^I$ where the two boundaries are situated in the undisturbed case. In order to find the amplitude z of the vertical motion at other levels, an additional integration has to be performed. If again quantities of higher than the first order in the perturbation terms are neglected,

$$\frac{dz}{dt} = \frac{\partial z}{\partial t} + U\frac{\partial z}{\partial x} = w\,.$$

By integration of this equation, for the upper layer it follows that

$$z^I = \left(\frac{g}{\alpha c^{I^2}}\sinh \alpha z + \cosh \alpha z\right) Z_0\, e^{i(\alpha x - \beta t)}\;,$$

and for the lower layer, if the expression (11) is used for K^{II},

$$(14)\qquad z^{II} = \frac{\sinh \alpha\,(z + H)}{\sinh \alpha h^{II}}\, Z_1\, e^{i(\alpha x-\beta t)}\;.$$

At the internal boundary, $z = -h^I$, the vertical displacements in both layers must be the same, which leads again to the relation (12). If by means of this relation Z_0 in the foregoing expression for z^I is replaced by Z_1,

$$(15)\qquad z^I = \frac{(g/\alpha c^{I2})\sinh \alpha z + \cosh \alpha z}{-(g/\alpha c^{I2})\sinh \alpha h^I + \cosh \alpha h^I}\, Z_1\, e^{i(\alpha x-\beta t)}\;.$$

As (14) shows, the amplitude of the vertical motion increases from zero at the bottom, $z = -H$, to Z_1 at the internal boundary, $z = -h^I$. In the upper layer the amplitude of the vertical motion at a given depth depends on the magnitude of c^I in addition to the parameters which determine the vertical amplitude in the lower layer. It is necessary, therefore, first to determine the wave velocity c from (13).

A rigorous solution is quite complicated because equation (13) is of the fourth degree, but approximate solutions can be obtained without much difficulty. For this purpose the conditions in the absence of currents, $U^I = U^{II} = 0$, may first be reviewed briefly. Let $\varepsilon = (\rho^{II} - \rho^I)/\rho^I$, which in all actual cases is small compared to unity. If powers higher than the first of ε are neglected, the following two well known roots of (13) are obtained:

$$(16)\quad c_1^2 = \frac{g}{\alpha}\,\frac{a^I + a^{II}}{1 + a^I a^{II}}\left[1 + \varepsilon\,\frac{a^{I2} + a^{II2} - a^{I2}a^{II2} - 1}{(a^I + a^{II})^2\,(1 + a^I a^{II})}\right]\quad\text{(surface wave)}$$

and

$$(17)\qquad c_2^2 = \frac{g}{\alpha}\,\frac{\varepsilon}{a^I + a^{II}}\quad\text{(internal wave)}.$$

The bracket in (16) is very closely equal to one and may be replaced by this value. The expressions for c_1^2 and c_2^2 show that, of the four roots of (13), the two for the internal waves are considerably smaller than the other two, at least as long as the density discontinuity is small, a condition which is always true for internal boundaries in the ocean.

For this reason it may now be assumed that in the more general case where U^I and U^{II} are not both zero, the values of c^I and c^{II} for the internal waves are also sufficiently small so that the first two terms in (13) may be neglected, these being of the fourth and third degree in these quantities. The solution of the remaining quadratic equation,

$$\text{(18)} \qquad \sigma \left(a^{II} c^{II^2} + a^I c^{I^2}\right) - \frac{g}{\alpha} (\sigma - 1) = 0 \, ,$$

results in the following expression for c,

$$\text{(19)} \quad c = \frac{U^I a^I + U^{II} a^{II}}{a^I + a^{II}} \pm \left[\frac{g}{\alpha} \frac{\rho^{II} - \rho^I}{\rho^{II}} \frac{1}{a^{II} + a^I} - \frac{(U^I - U^{II})^2}{(a^I + a^{II})^2} a^I a^{II} \right]^{\frac{1}{2}} .$$

In order to obtain a correction to this approximation let the rigorous solution of (13) be $c + j$, where j is a small quantity if c is a good approximation. If $c + j$ is substituted for c in (13), if powers higher than the first in j are neglected, and if it is taken into account that the value of c given by (18) makes the combination of the last two terms in (13) equal to zero, the following expression is obtained for the correction j to the approximate wave velocity

$$\text{(20)} \quad \left[4c^{I^3} + 2\sigma a^I a^{II} c^I c^{II} (c^I + c^{II}) - 2 \frac{g}{\alpha} \sigma \left(a^{II} c^{II} + a^I c^I\right) \right] j = - c^{I^4} - \sigma a^I a^{II} c^{I^2} c^{II^2} .$$

It should be noted that the value for c to be inserted in this formula is the one given by (19). For any particular numerical example it is simplest to check the approximation represented by (19) by direct computation of the correction term j from (20).

After the wave velocity c has been determined, the distribution of the vertical amplitude can be found from (14) and (15). As (14) shows, the velocity of the undisturbed current does not affect the distribution of the vertical amplitude in the lower layer if the amplitude of the internal discontinuity Z_1 is taken as the basic constant. But

U^I and U^{II} affect the value of the vertical amplitude in the upper layer, since (15) contains $c^I = c - U^I$, and since c depends on U^I and U^{II}. However, it can be seen that the effect of the current on the distribution of the vertical amplitude as a function of the depth will be small. With the actually occurring current velocities, c^I will be considerably smaller than g/α. Consequently, the first terms in the numerator and denominator of (15) will be considerably larger than the second terms, except very close to the free surface, $z = 0$. Consequently

$$z^I \approx -\frac{\sinh \alpha z}{\sinh \alpha h^I},$$

except in the closest vicinity of the free surface, and the distribution of the vertical amplitude is practically independent of the current velocity.

In order to amplify this statement, consider the smallest possible value of g/α for which the waves do not become exponentially unstable. In this case the square root in (19) is zero, so that

$$\frac{g}{\alpha} = \frac{(U^I - U^{II})^2}{a^I + a^{II}} a^I a^{II} \frac{\rho^{II}}{\rho^{II} - \rho^I},$$

and

$$\frac{g}{\alpha c^{I2}} = \frac{a^I + a^{II}}{a^{II}} a^I \frac{\rho^{II}}{\rho^{II} - \rho^I}.$$

Since the percentual density difference $(\rho^{II} - \rho^I)/\rho^{II}$ is a small quantity, $g/\alpha c^{I2}$ could be comparable to one only if $(a^I + a^{II})a^I/a^{II}$ were small. But since a^I is a hyperbolic cotangent it cannot be less than one, and it follows that a^{II} would have to be negative, which is impossible.

The case of long waves may also be considered separately. When the wave length is large compared to the depth of the fluid layer, the hyperbolic cotangent may be replaced by the reciprocal of its argument. The expression (19) for the velocity of the internal wave then becomes

$$(21) \quad c = \frac{U^I h^{II} + U^{II} h^I}{H} \pm \left[g \frac{h^I h^{II}}{H} \frac{\rho^{II} - \rho^I}{\rho^{II}} - \frac{(U^I - U^{II})^2 h^I h^{II}}{H^2} \right]^{\frac{1}{2}}.$$

Further, since the hyperbolic sine may be replaced by its argument, from (14)

$$z^{II} = \frac{z + H}{h^{II}} Z_1 \, e^{i(\alpha x - \beta t)} , \tag{22}$$

and since the hyperbolic cosine may be replaced by one,

$$z^{I} = \frac{\dfrac{gz}{c^{I^2}} + 1}{-\dfrac{gh^{I}}{c^{I^2}} + 1} Z_1 \, e^{i(\alpha x - \beta t)} . \tag{23}$$

Equation (21) shows that in the case of long waves the vertical amplitude increases linearly from the bottom to the internal boundary, as is well known in the absence of a basic current.

It can be shown similarly that in the upper layer the vertical amplitude decreases practically linearly from its maximum value at the internal boundary to the free surface. This statement is certainly true as long as the depth h^I of the upper layer is not very small, because for sufficiently large h^I the first terms in the numerator and denominator of (22) are much larger than one. Only for very small depths z the first term in the numerator of (22) will be of the order one, but at these levels the vertical motion due to the internal waves is very small anyway. If h^I is small compared to the total depth H of the fluid, it follows from (20) that approximately

$$c^{I^2} = gh^{I} \frac{\rho^{II} - \rho^{I}}{\rho^{II}} - \frac{(U^{I} - U^{II})^2 \, h^{I}}{H} .$$

If $h^I = 10$ m, $H = 1000$ m and $(\rho^{II} - \rho^{I})/\rho^{II} = 2 \cdot 10^{-3}$, for instance,

$$\frac{gh^{I}}{c^{I^2}} \sim \frac{\rho^{II}}{\rho^{II} - \rho^{I}} ,$$

so that the first terms in the denominator and numerator of (22) are still considerably larger than unity except very close to the surface, where the amplitude is very small, as pointed out before.

In order to show the modifications in the wave velocity produced by a current, the following example may be considered:

$\rho^{I} = 1.024$ $\qquad$ $\rho^{II} = 1.026$
$h^{I} = 100$ m $\qquad$ $h^{II} = 3000$ m
$L = 1000$ m.

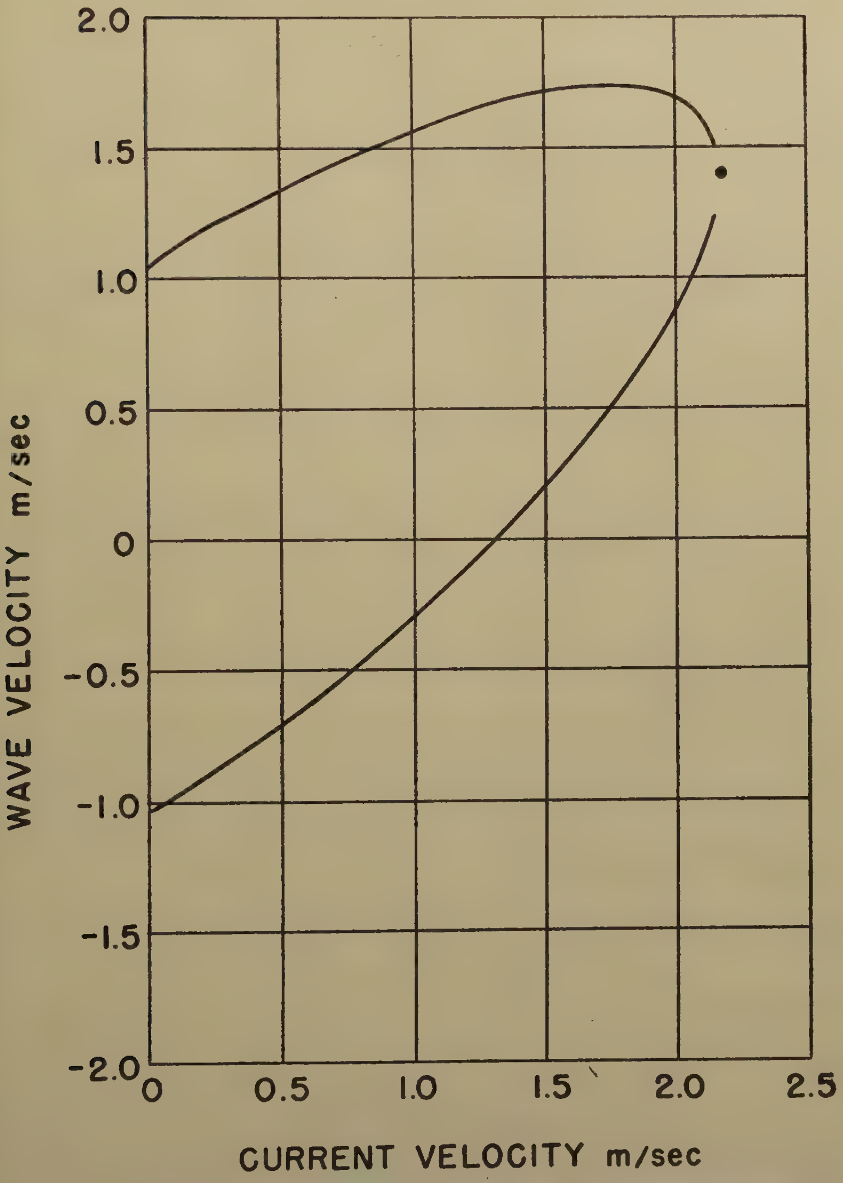


Figure 2. Wave velocity for different current velocities. (Density discontinuity 0.2%, depth of upper layer 100 m and of lower layer 3,000 m, wave length 1,000 m.)

The fluid in the lower layer may be at rest. In Fig. 2 the wave velocity is plotted as a function of the current velocity U^I in the upper layer. For each value of U^I two different values of c exist, depending on whether the positive or the negative sign before the square root is taken. The two solutions for c become identical when $U^I = 2.18$ m/sec, which represents the limit between stability and instability. The value of c for this value of U^I is indicated by a dot in Fig. 2, but the two curves are not joined here for the sake of clearness.

The depth distribution of the amplitude of the vertical motion can be computed from (14) and (15). The result is shown in Fig. 3 for the foregoing numerical example. As just explained, the velocity of the undisturbed current has only a very small effect on this distribution, and the curve shows consequently well known features, *viz.*, a rapid decrease from the internal boundary to the free surface in the relatively shallow surface layer and a less rapid decrease from the internal boundary to the bottom through the relatively deep lower layer.

As mentioned previously, only very close to the free surface does the amplitude of the vertical motion change substantially with different speeds of the undisturbed current. This fact is clearly shown by (12), which gives the amplitude ratio of the two boundaries. But even though the percentual variation of this ratio with different current velocities can be very large, its practical significance is very small, because the oscillation of the free surface produced by internal waves is negligible in any case.

Finally, the horizontal motion due to internal waves may be considered briefly. A certain modification of this motion occurs, of course, because it is now superimposed on basic currents, not on a state of rest. In addition, the amplitude of the horizontal motion will also be modified somewhat, as can be shown in the following manner. According to (9) and (12),

$$(24) \qquad u^I = \alpha c^I Z_1 \frac{(g/\alpha {c^I}^2)\cosh \alpha z + \sinh \alpha z}{-(g/\alpha {c^I}^2)\sinh \alpha h^I + \cosh \alpha h^I} e^{i(\alpha x - \beta t)}$$

and according to (3) and (11),

$$(25) \qquad u^{II} = \alpha c^{II} Z_1 \frac{\cosh \alpha (z + H)}{\sinh \alpha h^{II}} e^{i(\alpha x - \beta t)} \; .$$

In the numerator and denominator of (23) the second terms are, for most of the layer, much smaller than the first terms, as discussed before, so that u^I is largely proportional to c^I. Similarly, u^{II} is proportional to c^{II} according to (24). Any modifications in the value of

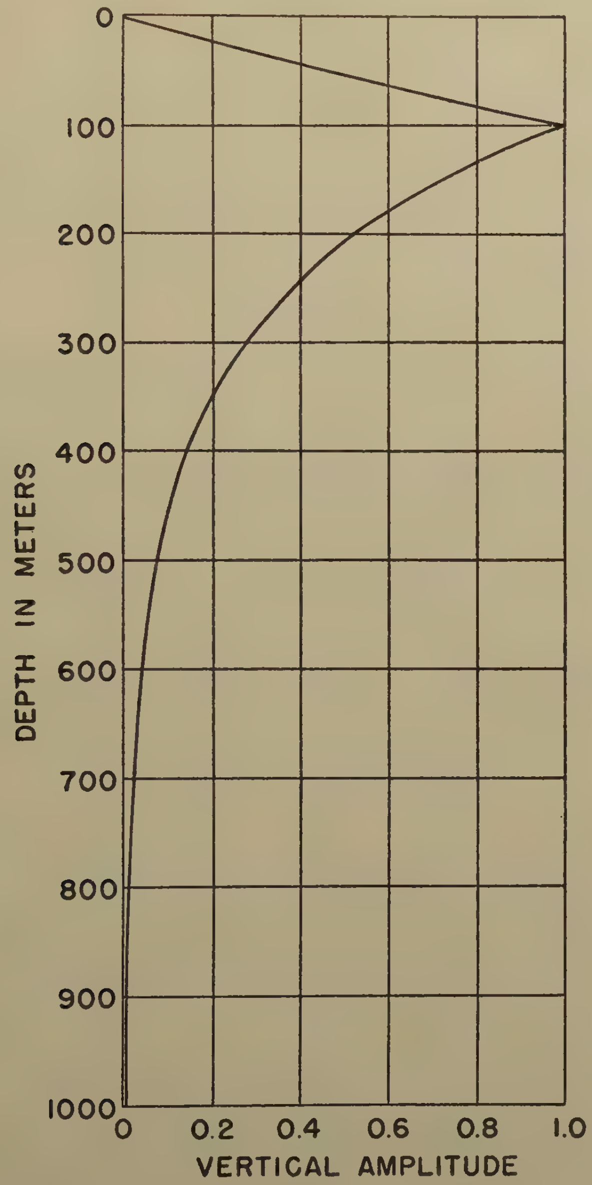


Figure 3. Distribution of the amplitude of the vertical motion as function of depth (amplitude at the internal discontinuity equal to one).

c caused by the existence of an undisturbed current will therefore be reflected in the amplitudes of u^{I} and u^{II}.

SUMMARY

The effect of a basic current on internal waves in the ocean is studied. It is shown that, while the wave velocity is modified by the existence of such a current, the distribution of the vertical amplitude of the wave motion as a function of the depth is very little affected by the presence of an ocean current. The amplitude of the horizontal motion is modified to a certain extent by the presence of ocean currents.

REFERENCES

Ekman, V. W.
1904. On dead water. Sci. Res. Norweg. N. Pol. Exped., 1893–1896, *5* (15): 1–152.
Seiwell, H. R.
1937. Short period vertical oscillations in the western basin of the North Atlantic. Pap. phys. Oceanogr. Meteorol., *5* (2): 1–44.
1942. An analysis of vertical oscillations in the southern North Atlantic. Proc. Amer. phil. Soc., *85* (2): 136–158.
Sverdrup, H. U., M. W. Johnson and R. H. Fleming.
1942. The oceans. Prentice-Hall, Inc., New York. 1087 pp.
Ufford, C. W.
1947. Internal waves in the ocean. Trans. Amer. geophys. Un., *28* (1): 79–86.
1947. Internal waves measured at three stations. Trans. Amer. geophys. Un., *28* (1): 87–95.

THE TRANSMISSION OF ENERGY BY GRAVITY WAVES OF FINITE HEIGHT

By

VICTOR P. STARR

Massachusetts Institute of Technology

AND

GEORGE W. PLATZMAN

University of Chicago

1. *Introduction.* One of the obvious properties of a train of gravity waves is the transmission of energy in the direction of wave propagation. It is therefore a matter of considerable interest to compute the rate of this energy transfer for waves of given characteristics. This problem has been studied by Reynolds (1877) and by Rayleigh (1877) for the case of irrotational waves, and their results are discussed in standard textbooks on hydrodynamics such as the one by Lamb (1932). It remains to be pointed out, however, that these classical treatments of the subject involve the simplifying assumption that the waves are of infinitesimal amplitude. To the knowledge of the writers, there does not exist any treatment which deals with the question in its generality when this simplifying assumption is dropped. The purpose of this paper is to supply such a treatment. Certain important consequences of the results obtained will form the subject matter of a further discussion by the writers at a later time.

2. *The Energy Transmission Function.* Let us consider the two-dimensional irrotational motions in the x, z-plane of a perfect fluid of uniform and constant density ρ, bounded above by a free surface at which the pressure p has a uniform value p_0 and below by a fixed bottom of given shape, as shown in Fig. 1. To the right (x positive) the fluid is bounded by a fixed vertical wall at $x = x_2$, while to the left the fluid may extend to great distances. It is assumed that there exists an acceleration of gravity g acting downward, *i. e.*, in the negative z-direction. The origin for z is left unspecified for the time being. Also, for the present purpose no additional assumptions will be made concerning the detailed nature of the fluid motion except that the free surface remains unbroken so that there is no turbulent dissipation of energy. We shall now attempt to deduce an expression for the rate of energy transfer in the fluid across an arbitrary fixed vertical located at $x = x_1$.

Generally speaking, the energy flux across $x = x_1$ in the fluid will, if energy is conserved, be equal to the rate of increase of energy in the fluid to the right plus the rate at which work is performed against the surface pressure in the region to the right. Therefore, terms corresponding to these processes should appear in the final equation obtained.

Since the motion is irrotational, there exists a velocity potential ϕ so that the particle velocity components u and w in the x- and z-directions respectively are given by

$$u = -\frac{\partial \phi}{\partial x}; \quad w = -\frac{\partial \phi}{\partial z}. \tag{1}$$

We also have at our disposal the continuity equation,

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0. \tag{2}$$

According to the theory of the motions of an ideal fluid, as given for example by Lamb (1932), it is possible to write an integral of the dynamic equations of motion in the form

$$\tfrac{1}{2} \rho q^2 + \rho g z + p = \rho \frac{\partial \phi}{\partial t} + F(t). \tag{3}$$

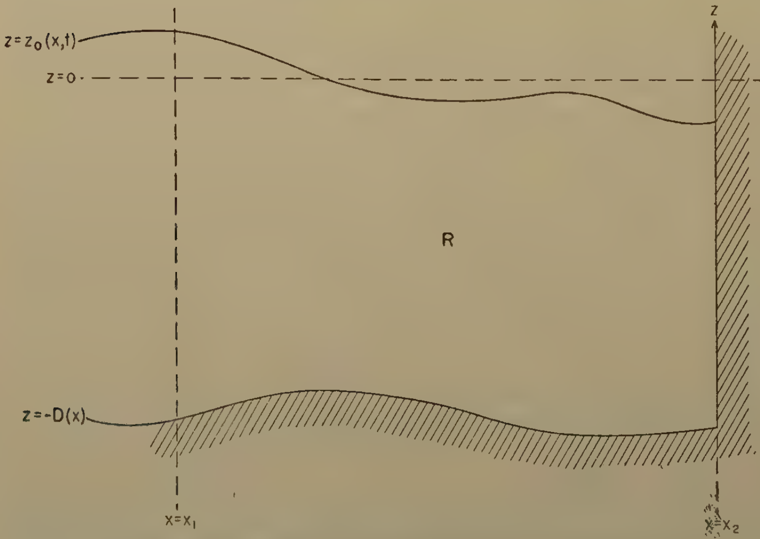


Figure 1. Schematic diagram for the computation of the energy transfer due to a fluid layer across the vertical at $x = x_1$.

Here $q^2 \equiv u^2 + w^2$, t is time, and $F(t)$ is an arbitrary function of time alone. Indicating the expression on the left-hand side of (3) by

$$\chi \equiv \tfrac{1}{2}\rho q^2 + \rho g z + p\,,$$

we may multiply by $(udz - wdx)$ and write

$$\chi\,(udz - wdx) = \left[\rho\frac{\partial\phi}{\partial t} + F(t)\right](udz - wdx)\,. \tag{4}$$

The quantity on the left represents the instantaneous flux of χ through an elementary segment of the boundary of an arbitrary region of the fluid. Let us now form the line integral of (4) in the counterclockwise sense around the boundary of the region R (Fig. 1). It is apparent that the contribution from $F(t)$ will vanish by virtue of (2), while the remaining terms on the right-hand side may be expressed as a surface integral with the aid of Stokes' theorem and be immediately simplified by use of (1) and (2). Thus we may state that

$$\begin{aligned}\oint\chi\,(udz - wdx) &= \iint\limits_R\left[\frac{\partial}{\partial x}\left(\rho u\frac{\partial\phi}{\partial t}\right) + \frac{\partial}{\partial z}\left(\rho w\frac{\partial\phi}{\partial t}\right)\right]dx\,dz\\ &= -\iint\limits_R\frac{\partial}{\partial t}\left(\tfrac{1}{2}\rho q^2\right)dx\,dz\,.\end{aligned}\tag{5}$$

Remembering that no contributions to the line integral on the left can result from the portions of the circuit represented by the bottom and the wall, and also remembering that the position of the upper boundary is a function of time, we may rewrite (5) as follows:

$$\begin{aligned}&\int_{-D_1}^{z_{01}}\chi u dz + \int_{x_1}^{x_2}\chi_0\,(u_0 dz_0 - w_0 dx)\\ &= \frac{\partial}{\partial t}\iint\limits_R\tfrac{1}{2}\rho q^2\,dx\,dz - \int_{x_1}^{x_2}\tfrac{1}{2}\rho q_0{}^2\frac{\partial z_0}{\partial t}\,dx\,.\end{aligned}\tag{6}$$

The zero subscripts are used to indicate values of the variables at the free surface, while the subscripts 1 and 2 have the significance indicated in Fig. 1.

In view of the fact that the free surface is a material surface through which no fluid passes, we may write that (see Starr, 1945)

$$w_0 = \frac{\partial z_0}{\partial t} + u_0\frac{\partial z_0}{\partial x}\,; \quad \frac{\partial z_0}{\partial x}dx = dz_0\,. \tag{7}$$

The first of these is the kinematic boundary condition for the free surface, the dynamic boundary condition being furnished by the uniformity of pressure. From the definition of χ and by the use of relations (7), it follows that (6) may be put in the form

$$\int_{-D_1}^{z_{01}} (\tfrac{1}{2}\rho q^2 + \rho g z + p)\, u dz$$
$$= \frac{\partial}{\partial t} \iint_R \tfrac{1}{2}\rho q^2 \, dx\, dz + \frac{\partial}{\partial t} \int_{x_1}^{x_2} \tfrac{1}{2}\rho g z_0^2 \, dx + p_0 \int_{x_1}^{x_2} \frac{\partial z_0}{\partial t} dx \tag{8}$$

The first term on the right-hand side of (8) represents the rate of increase of kinetic energy of the fluid to the right of $x = x_1$. The second term is the rate of increase in the potential energy of the same fluid, while the last term is the rate at which energy is lost from the region R through work done against the surface pressure p_0. From the principle of conservation of energy we infer that the expression on the left-hand side of (8) must give the rate of energy transfer within the fluid across the fixed vertical at $x = x_1$. This expression, therefore, we call the *energy transmission function.*

Before proceeding with applications to wave phenomena, several additional comments and observations must be made in regard to equation (8). By continuity considerations one may write that

$$\int_{x_1}^{x_2} \frac{\partial z_0}{\partial t} dx = \int_{-D_1}^{z_{01}} u dz \, . \tag{9}$$

It then follows from (8) that the *rate of transmission of energy utilized in changing the potential and kinetic energy of the fluid to the right of* $x = x_1$, is given by

$$\int_{-D_1}^{z_{01}} [\tfrac{1}{2}\rho q^2 + \rho g z + (p - p_0)]\, u dz \, . \tag{10}$$

The quantity $p - p_0$ is the portion of the pressure *due to the fluid,* and is equal to the actual pressure when p_0 is zero. Since the portion of the total pressure represented by p_0 is spacially uniform, it cannot change the motions of the fluid, and no generality is lost by assuming it to be zero, as will be done henceforth.

In equation (8) both the potential energy in the region and its variation with time are not independent of the choice of a reference level, namely the origin for z. It follows, therefore, that the transmission of energy should also be indeterminate to this extent, as is indeed the case, since z occurs in the integrand on the left-hand side of (8). Although formally the choice for the origin of z is arbitrary, we shall follow the customary convention used in wave studies and place it at the undisturbed level of the fluid, that is, at the mean level of the free surface.

If the fluid extends vertically downward to very great depths where the particle motions disappear, no generality is lost by extending the integration to $z = -\infty$ in the transmission function.

3. *Periodic Waves in a Deep Basin.* Using the several conventions decided upon in the preceding section, we next turn our attention to the special circumstances in the case of periodic surface waves in a deep fluid, where the motions disappear for great depths.[1] We desire to find the energy transferred across a fixed vertical during one wave period T. Accordingly, from the transmission function, this amount of energy e_t is

$$e_t = \int_0^T \int_{-\infty}^{z_0} \chi u dz dt \,. \tag{11}$$

Since we assume that the waves move without change in their structure, it is possible to replace the time integration in (11) by a space integration over one wave length. Thus we employ the transformation

$$dx = cdt \,; \quad L = ct \,, \tag{12}$$

where c is the wave speed and L is the wave length, so that (11) becomes

$$e_t = 1/c \int_{-\infty}^{z_0} \int_0^L \chi u dx dz \,. \tag{13}$$

Assuming the waves to be moving to the right in Fig. 1, and denoting the particle velocity in the x-direction relative to the moving waves by U, so that

$$U = u - c \,, \tag{14}$$

it can be shown (see Lamb, 1932; Starr, 1947) that in the present case

---

[1] In view of the fact that the exact solution for periodic waves implies the existence of a wave train of infinite length, it might appear that the argument presented in the previous section cannot be applied. This apparent difficulty can be avoided by supposing that the distance between x_1 and x_2 in Fig. 1 is very large and is occupied by a forerunner of great length, such that the waves present at $x = x_1$ are for practical purposes of the same structure as in an ideal infinite train. The vertical wall at $x = x_2$ would then be equivalent to a region of resting fluid ahead of the long forerunner.

$$\tfrac{1}{2}\rho\,(U^2 + w^2) + \rho g z + p = \tfrac{1}{2}\rho c^2\,, \tag{15}$$

which is simply the Bernoulli equation for the (steady) relative motion. From (14) and (15) it follows that $\chi = \rho c u$, and consequently, according to (13), *the energy transmitted during one wave period is equal to twice the kinetic energy of the horizontal motions per wave length*, e_x. Thus

$$e_t = \int_{-\infty}^{z_0}\int_0^L \rho u^2 dx dz = 2\,e_x\,. \tag{16}$$

In order to examine the import of equation (16), we next form the ratio of the energy transmitted during a wave period to the total wave energy per wave length. This total wave energy is the sum of the total kinetic energy e and the wave potential energy v, the latter quantity being defined as the excess of potential energy due to the presence of the wave disturbance, *i. e.*, $v = \int_0^L \tfrac{1}{2}\rho g z_0^2 dx$. The desired ratio is then

$$\frac{e_t}{e+v} = \frac{2\,e_x}{e+v}\,. \tag{17}$$

In a previous communication (Starr, 1947), it was shown that under the present circumstances, if e_z is the kinetic energy of the vertical motions per wave length so that $e_x + e_z = e$, then

$$e_x - e_z = 2\,(e - v)\,, \tag{18}$$

without approximation. With the aid of (18), (17) may be put in the form

$$\frac{e_t}{e+v} = \tfrac{1}{2} + \frac{5}{2}\,\frac{e-v}{e+v}\,. \tag{19}$$

According to results previously obtained (Platzman, 1947), e exceeds v except for infinitesimal amplitudes when equality of the two quantities is approached. Equation (19) is thus in agreement with the classical small-amplitude theory, according to which the ratio of transmitted energy to energy 'density'—and correspondingly the ratio of group velocity[2] to phase velocity—is equal to one half.

---

[2] More appropriately, *dynamical* group velocity. A better term would probably be *energy velocity.*

But it is apparent that for waves of *finite* height the above-mentioned ratio is *larger* than one half. For waves of the extreme form it was estimated in the reference mentioned that

$$e - v \approx \tfrac{1}{8} v . \tag{20}$$

This leads to a value of about 11/17 for the maximum value of the ratio, *i. e.*,

$$\frac{e_t}{e + v} \approx \frac{11}{17} \approx 0.65 ,$$

for waves of maximum height.

4. *Partition of the Energy Transport for Periodic Waves.* It is a matter of some interest to examine the individual contributions of the three terms in the transmission function (10) to the energy transfer per wave period as given by (16) for periodic waves in a deep basin. For this purpose we may conveniently employ the symbolism

$$\{H\} \equiv 1/c \int_{-\infty}^{z_0} \int_{0}^{L} H\, u\, dx\, dz ,$$

in which all quantities have the meaning already attached to them, and H is any function of x and z. On reference to (13) it may be seen, for example, that

$$e_t = \{\chi\} = \{\tfrac{1}{2} \rho q^2\} + \{\rho g z\} + \{p_0\} .$$

The transmitted energy e_t may thus be taken as the sum of three quantities:—

a. The *advection of kinetic energy:*

$$\{\tfrac{1}{2} \rho q^2\} \equiv 1/c \int_{-\infty}^{z_0} \int_{0}^{L} \tfrac{1}{2} \rho q^2\, u\, dx\, dz .$$

In order to evaluate this quantity, let us make the substitution $u = c + U$, so that

$$\{\tfrac{1}{2} \rho q^2\} = e + 1/c \int_{-\infty}^{z_0} \int_{0}^{L} \tfrac{1}{2} \rho q^2\, U\, dx\, dz , \tag{21}$$

where e is the total kinetic energy, as employed in the preceding section. In the same way we find that

$$\{\rho u\} = m + 1/c \int_{-\infty}^{z_0} \int_{0}^{L} \rho u\, U\, dx\, dz , \tag{22}$$

where

$$m \equiv \int_{-\infty}^{z_0}\int_{0}^{L} \rho u \, dx \, dz \,,$$

is the wave momentum. On taking account of the relation $mc = 2e$, established previously (Starr, 1947), and of the identity $\{\rho u\} \equiv 2e_x/c$ where e_x is the kinetic energy of the horizontal wave motions, equation (22) may be written

$$0 = 2\, e_z + 1/c \int_{-\infty}^{z_0}\int_{0}^{L} \rho u c \, U \, dx \, dz \,. \tag{23}$$

Subtracting (23) from (21),

$$\{\tfrac{1}{2}\, \rho q^2\} = (e_x - e_z) + 1/c \int_{-\infty}^{z_0}\int_{0}^{L} (\tfrac{1}{2}\, \rho q^2 - \rho u c)\, U \, dx \, dz \,.$$

In this last equation the integral on the right vanishes identically,[3] so that finally

$$\{\tfrac{1}{2}\, \rho q^2\} = e_x - e_z = 2\,(e - v)\,, \tag{24}$$

the latter equality being given by (18). *The amount of kinetic energy transported by the wave motion in one wave period is equal to the difference between the horizontal and vertical kinetic energies per wave length.*

b. The *work done by pressure forces:*

$$\{p\} \equiv 1/c \int_{-\infty}^{z_0}\int_{0}^{L} p u \, dx \, dz \,.$$

Consider the identity

$$pu = -\, p \frac{\partial \phi}{\partial x} = -\, \frac{\partial p \phi}{\partial x} + \phi \frac{\partial p}{\partial x} \,.$$

In the last term we may make the substitution $\partial p/\partial x = -\, \rho\, du/dt$, which is merely the dynamical equation of motion for the x-direction. We obtain, after one further transformation,

$$pu = -\, \frac{\partial p \phi}{\partial x} - \rho \frac{d \phi u}{dt} + \rho u \frac{d \phi}{dt} \,. \tag{25}$$

---

[3] To demonstrate this, we may replace $U dz$ by $-d\Psi$, where Ψ is the stream function for the motion relative to the moving wave. The integration may then be conceived as taking place first with respect to x along the relative streamlines Ψ = constant and then vertically from the free surface to great depths. But in the first stage of this process the integrand $\tfrac{1}{2}\rho q^2 - \rho u c$, which according to (14) and (15) may be replaced by $-\rho g z - p$, will vanish along each relative streamline, since the pressure along these lines must be *on the average* hydrostatic (see Starr, 1947).

On performing the integration indicated in the expression for $\{p\}$, we find that the contribution of the first term on the right in (25) vanishes identically because of the periodicity of $p\phi$ and because of the vanishing of the surface pressure. Likewise, the contribution of the second term on the right in (25) vanishes because of the periodicity of ϕu and because of the kinematical boundary condition requiring the free surface to be a streamline in the motion relative to the wave.[4] We see, therefore, that the last term on the right in (25) expresses the average work of pressure forces:

$$\{p\} = \{\rho\, d\phi/dt\}\,.$$

Now

$$\frac{d\phi}{dt} = \frac{\partial\phi}{\partial t} + u\frac{\partial\phi}{\partial x} + w\frac{\partial\phi}{\partial z} = cu - q^2\,,$$

hence

$$\{p\} = c\,\{\rho u\} - \{\rho q^2\}\,.$$

But $c\,\{\rho u\} = 2\,e_x$ and, from (24), $\{\rho q^2\} = 2\,(e_x - e_z)$. Therefore,

$$\{p\} = 2e_z\,. \tag{26}$$

The work done by pressure forces operating in one wave period is equal to twice the vertical kinetic energy per wave length.

c. The *advection of potential energy:*

$$\{\rho g z\} \equiv 1/c \int_{-\infty}^{z_0}\int_{0}^{L} \rho g z\, u\, dx\, dz\,.$$

From equations (13), (24), and (26),

$$\begin{aligned}\{\rho g z\} &= \{\chi\} - \{\tfrac{1}{2}\rho q^2\} - \{p\}\\ &= (2\,e_x) - (e_x - e_z) - (2\,e_z)\,.\end{aligned}$$

Hence

$$\{\rho g z\} = e_x - e_z = \{\tfrac{1}{2}\rho q^2\}\,, \tag{27}$$

---

[4] To demonstrate this, note that

$$\frac{d\phi u}{dt} = U\frac{\partial\phi u}{\partial x} + w\frac{\partial\phi u}{\partial z}\,,$$

or, in virtue of the continuity equation,

$$\frac{d\phi u}{dt} = \frac{\partial U\phi u}{\partial x} + \frac{\partial w\phi u}{\partial z}\,.$$

The space integration may now be replaced by a line integral of $\phi u\,(U dz - w dx)$ taken around the boundary of the region of integration, and the stated result follows directly.

the latter equality being given by (24). *The amount of potential energy transported by the wave motion in one wave period is equal to the amount of kinetic energy transported in the same period.*

SUMMARY

In this paper an expression is derived for the transfer of energy across a vertical plane, due to the arbitrary two-dimensional motions of a layer of ideal fluid under gravity. This expression—the energy transmission function—is comprised of three terms, representing the advection of kinetic and of potential energy together with the work done by pressure forces.

The energy transmission function is applied to the problem of energy transfer produced by gravity waves of finite height in a deep basin. The result obtained indicates that the energy transferred per wave period is greater than one-half of the total energy per wave length.

REFERENCES

LAMB, HORACE

1932. Hydrodynamics. Cambridge Univ. Press, sixth ed. 738 pp.

PLATZMAN, G. W.

1947. The partition of energy in periodic irrotational waves on the surface of deep water. J. Mar. Res., *6* (3): 194–202.

RAYLEIGH, LORD

1877. On progressive waves. Proc. Lond. math. Soc., (1) *9:* 21–26. (Also in Rayleigh's *Scientific Papers*, *1*, 1899: 322–327. Cambridge Univ. Press.)

REYNOLDS, OSBORNE

1877. On the rate of progression of groups of waves, and the rate at which energy is transmitted by waves. *Nature*, Lond., *16:* 343–344. (Also in Reynolds' *Scientific Papers*, *1*, 1900: 198–203. Cambridge Univ. Press.)

STARR, V. P.

1945. A quasi-Lagrangian system of hydrodynamical equations. J. Meteor., *2* (4): 227–237.

1947. Momentum and energy integrals for gravity waves of finite height. J. Mar. Res., *6* (3): 175–193.

REFLECTION OF TSUNAMIS[1]

By

J. D. COCHRANE AND R. S. ARTHUR[2]

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

Waves are reflected whenever they move into a region where their speed of propagation changes. It has generally been found that for appreciable reflection a large change in speed of propagation must occur over a distance which is small compared to a wave length. Since tsunamis are shallow water waves, they are propagated at a speed c which depends on the depth h according to

$$c = \sqrt{gh} .$$

A typical tsunami, one having a period of 20 minutes for example, has a wave length of 90 miles in water of 5,000-ft. depth. At the continental slopes and walls of deep oceanic trenches, an appreciable depth change takes place over a distance which is small compared with this typical wave length, and reflection is to be expected. The object of this paper is to examine the observational evidence for tsunami reflection and to make a comparison with results which are derived theoretically.

OBSERVATIONAL EVIDENCE OF REFLECTION

Occasionally a tide gauge record of a tsunami shows a marked amplitude increase long after the arrival of the first waves. Such a feature has been considered as possible evidence of reflection if its time of occurrence coincides with the expected arrival time of waves traveling

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 377. This work represents results of research carried out for the Office of Naval Research and the Hydrographic Office of the Department of the Navy under contract with the University of California.

[2] We would like to take this occasion to express our great appreciation for the privilege of working and studying under the direction of Dr. H. U. Sverdrup. The training and guidance which he has provided have been and will continue to be invaluable to us.

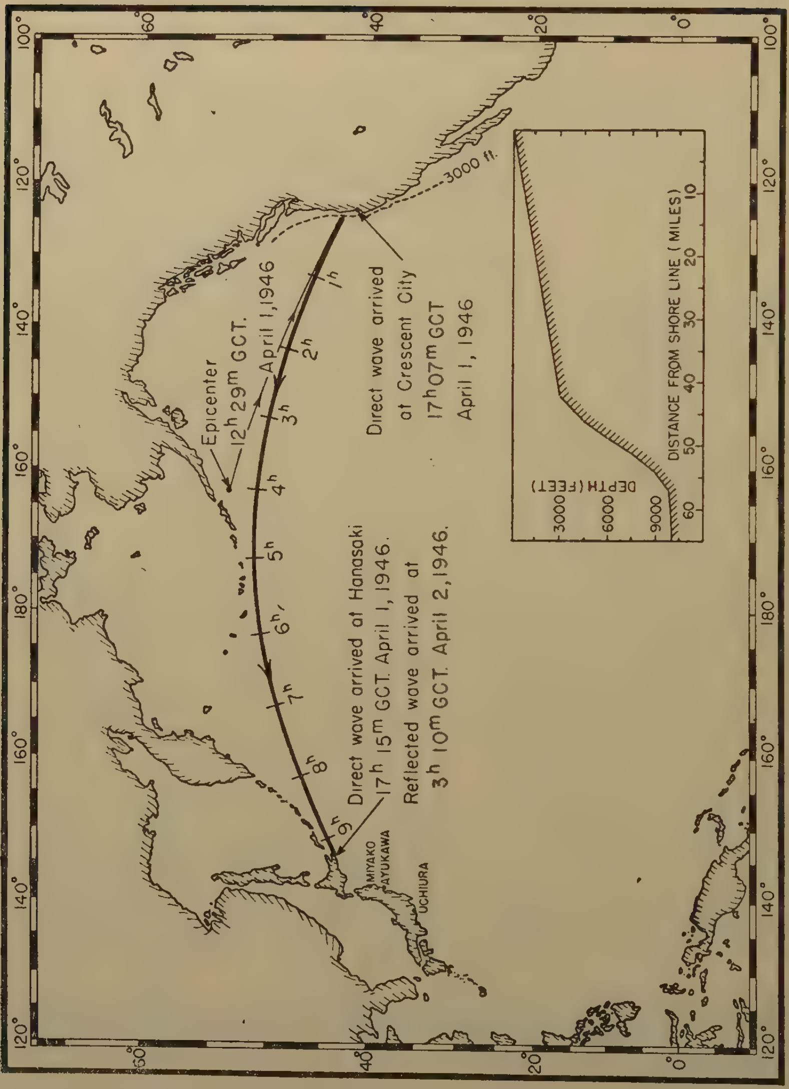

Epicenter
12h29m GCT.
April 1,1946
1h
2h
3h
4h
5h
6h
7h
8h
9h
Direct wave arrived
at Crescent City
17h07m GCT
April 1, 1946
3000 ft.
Direct wave arrived at Hanasaki
17h 15m GCT. April 1, 1946.
Reflected wave arrived at
3h 10m GCT. April 2,1946.
MIYAKO
AYUKAWA
UCHIURA
DEPTH (FEET)
3000
6000
9000
DISTANCE FROM SHORE LINE (MILES)
10
20
30
40
50
60
100°
120°
140°
160°
180°
160°
140°
120°
60°
40°
20°
0°

Figure 1. Arrival times at Hanasaki of direct and reflected waves for tsunami of April 1, 1946. Heavy solid line indicates probable path of reflected waves reaching Japanese Coast. Inset shows a typical bottom profile along a line from shore near Crescent City.

over a path involving at least one reflection. Amplitude variations appear frequently in records of tsunamis. The evidence for reflection is most conclusive, therefore, when the amplitude increases greatly during a short interval of time and then remains relatively large for a period of several hours.

Two records discussed by Hart (1931: 192–198) are ideal in this respect, since the tsunamis were associated with earthquakes east of New Zealand, and the records show no indication of tsunami-type waves at Sydney until the arrival of waves supposedly reflected. Of about 25 additional records examined by the present writers, those from four Japanese stations (location in Fig. 1) for the tsunami of April 1, 1946 contain indication of reflection[3] (Figs. 2–5), and of these only the evidence from Hanasaki is conclusive.

To bring out the general trend of wave height, individual wave heights were averaged in overlapping two-hour intervals centered at each of the first 20 hours of the tsunami. The arrival of a train of reflected waves may be expected to show up as a relatively steep slope in the curve of average height.

The first few waves were small in all four cases, hardly exceeding 10 cm. at most. After seven to ten hours had passed, the average heights had increased two- to seven-fold (Figs. 2–5). The most sudden increase occurred at Hanasaki, where the average height grew from 6 cm. to 24 cm. in the time interval 0200 to 0500 G. C. T. The slope of the curve of average height is steepest at 0310 G. C. T., 9 hr. and 55 min. after the arrival of the first wave. A rough refraction diagram indicates that waves could reach Hanasaki by way of a reflection at or near the coast of southern Oregon (Fig. 1). The expected time interval from the arrival of the first wave to the arrival of such a reflected wave is 9 hr. and 24 min., which is close to the interval actually observed.

THEORY

The reflecting mechanisms which are considered in this paper are the rather abrupt changes in depth embodied in the continental slopes and walls of deep oceanic trenches. As a first approximation, the size of reflected waves is calculated for a simple model defined as follows:

1) The depth changes abruptly along a straight line (Fig. 6, showing the co-ordinate system which is used).

---

[3] Evidence from records and a refraction diagram received from the Japanese Central Meteorological Observatory through the Hydrographic Office, Department of the Navy. In the refraction diagram prepared by the observatory, wave trains reflected near various sections of the North American coast are shown.

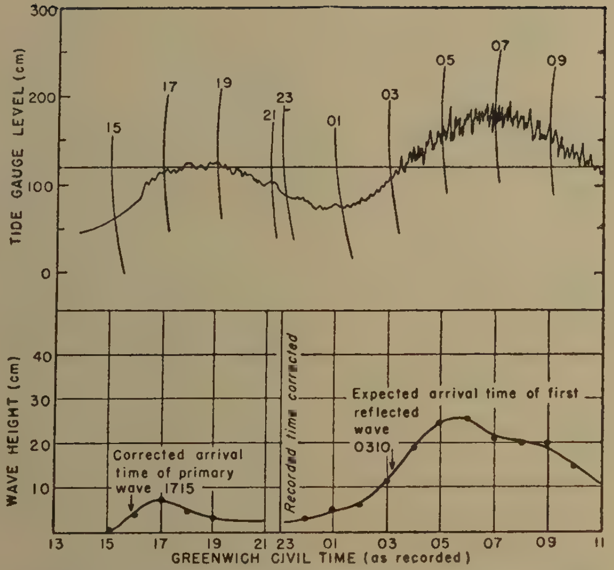


FIGURE 2. Hanasaki tide gauge record for tsunami of April 1, 1946, and graph of average wave height.

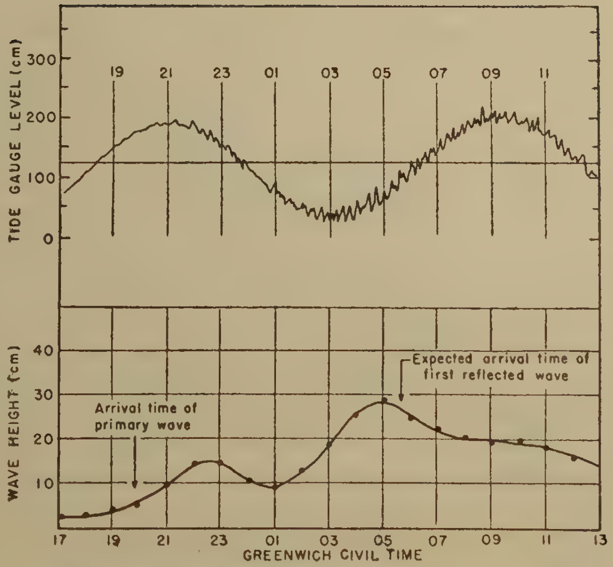


FIGURE 3. Miyako tide gauge record for tsunami of April 1, 1946, and graph of average wave height.

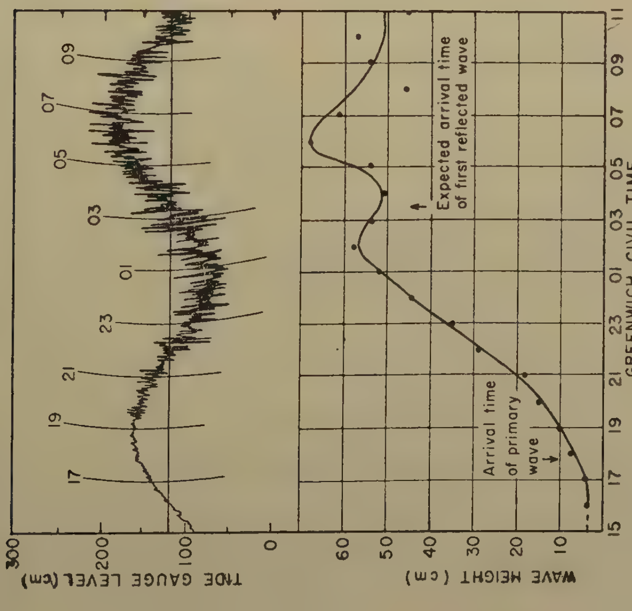


Figure 4. Ayukawa tide gauge record for tsunami of April 1, 1946, and graph of average wave height.

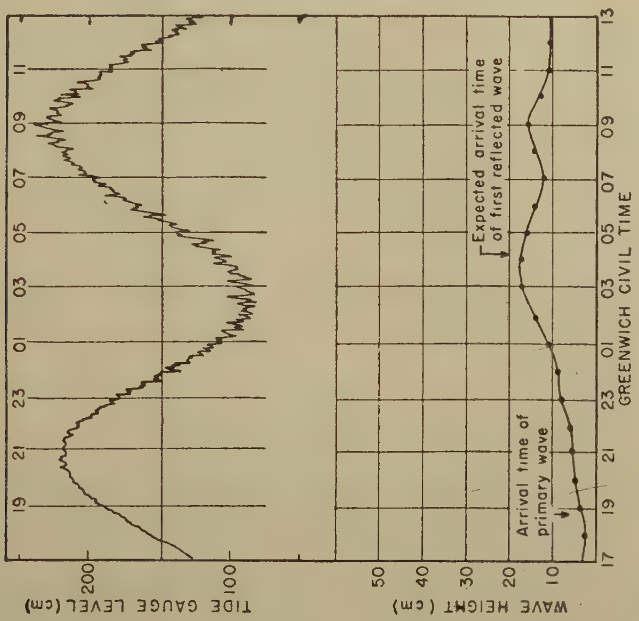


Figure 5. Uchiura tide gauge record for tsunami of April 1, 1946, and graph of average wave height.

2) The waves striking this discontinuity in depth are plane and sinusoidal.
3) The ratio of the depth h to the wave length L is less than 0.05, so that the theory of "shallow water" waves holds.
4) The amplitude of the displacement of the free surface, ζ, is very small compared with L and h.

The differential equation of waves satisfying conditions (3) and (4) of the model is for $x > 0$ (Fig. 7)

$$\frac{\partial^2 \zeta_1}{\partial t^2} = c_1^2 \left(\frac{\partial^2 \zeta_1}{\partial x^2} + \frac{\partial^2 \zeta_1}{\partial y^2} \right), \quad (c_1 = \sqrt{gh_1}) \, ,$$

and for $x < 0$,

$$\frac{\partial^2 \zeta_2}{\partial t^2} = c_2^2 \left(\frac{\partial^2 \zeta_2}{\partial x^2} + \frac{\partial^2 \zeta_2}{\partial y^2} \right), \quad (c_2 = \sqrt{gh_2}) \, .$$

Solutions providing for plane sinusoidal incident and reflected waves on the negative-x side, and transmitted waves on the positive-x side, are:

$$\zeta_1 = A_1 \cos \sigma \left(t + \frac{\mu_1 x + \nu_1 y}{c_1} \right) + A_1' \cos \sigma \left(t + \frac{\mu_1' x + \nu_1' y}{c_1} \right), \tag{1}$$

$$\zeta_2 = A_2 \cos \sigma \left(t + \frac{\mu_2 x + \nu_2 y}{c_2} \right), \tag{2}$$

where $\sigma = 2\pi/T$, T being the period,

and
$$\begin{aligned} \mu_1 &= \cos \theta_1, & \nu_1 &= \sin \theta_1, \\ \mu_1' &= \cos (\pi - \theta_1') = -\cos \theta_1', & \nu_1' &= \sin (\pi - \theta_1') = \sin \theta_1', \\ \mu_2 &= \cos \theta_2, & \nu_2 &= \sin \theta_2; \end{aligned}$$

these angles being shown in Fig. 7.

The displacement, ζ, and mass transport perpendicular to the discontinuity in depth, $\rho h u$, are continuous across the discontinuity. That is, at $x = 0$

$$\zeta_1 = \zeta_2, \tag{3}$$

and

$$\rho h_1 u_1 = \rho h_2 u_2. \tag{4}$$

Since u is proportional to $\partial \zeta / \partial x$ for sinusoidal waves, (4) may be expressed by

$$h_1 \frac{\partial \zeta_1}{\partial x} = h_2 \frac{\partial \zeta_2}{\partial x} \, . \tag{4'}$$

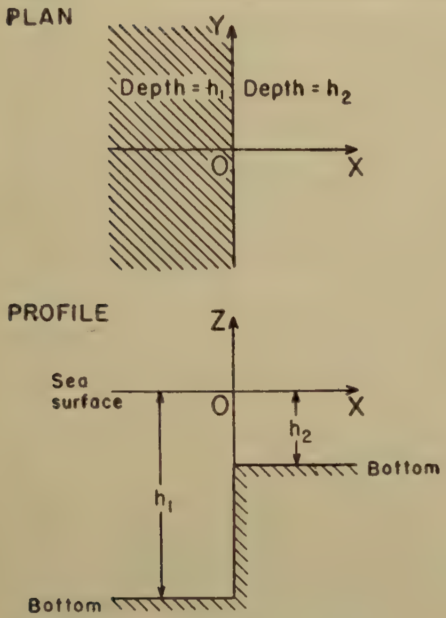

PLAN
Y
Depth = h_1
Depth = h_2
O
X
PROFILE
Z
Sea surface
O
X
h_2
Bottom
h_1
Bottom

FIGURE 6. Schematic diagram showing abrupt depth change assumed in theory.

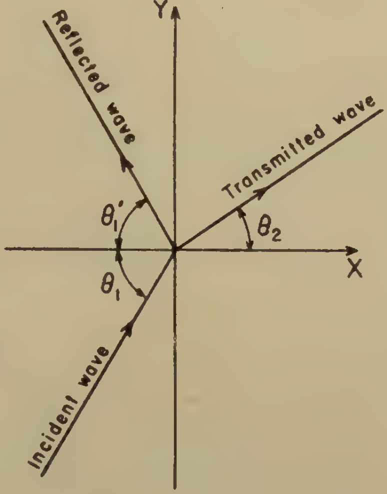

Y
Reflected wave
Transmitted wave
θ_1'
θ_2
X
θ_1
Incident wave

FIGURE 7. Direction of travel of incident wave, reflected wave, and transmitted wave.

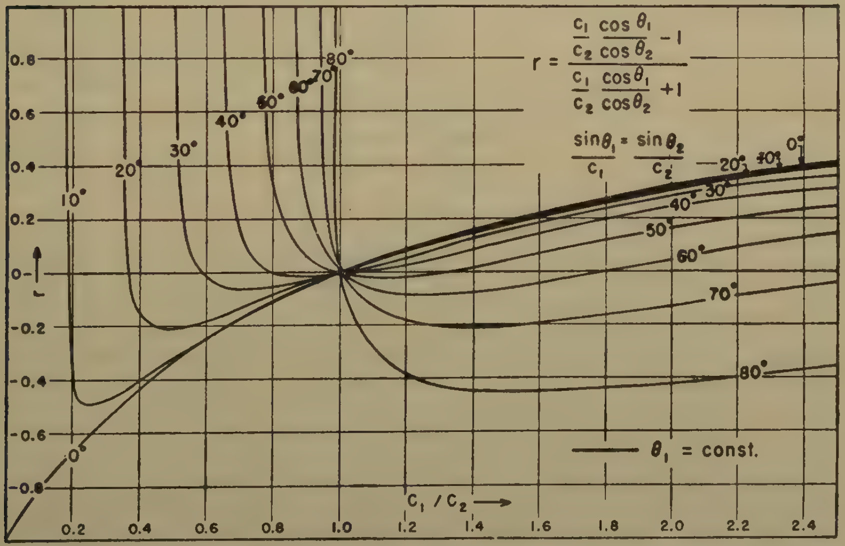


FIGURE 8. Amplitude coefficient of reflection, r, as a function of the velocity ratio, c_1/c_2, for various values of the angle of incidence θ_1.

Substituting (1) and (2) into (3) and (4′) gives

$$A_1 + A_1' = A_2, \tag{5}$$

and

$$\frac{h_1\mu_1A_1}{c_1} + \frac{h_1\mu_1'A_1'}{c_1} = \frac{h_2\mu_2A_2}{c_2}, \tag{6}$$

at $x = y = 0$. It follows from (5) and (6) that the ratios A_1'/A_1 and A_2/A_1 are given by

$$\frac{A_1'}{A_1} = \frac{c_2\mu_2 - c_1\mu_1}{c_1\mu_1' - c_2\mu_2}, \tag{7}$$

and

$$\frac{A_2}{A_1} = \frac{c_1\mu_1' - c_1\mu_1}{c_1\mu_1' - c_2\mu_2}. \tag{8}$$

The ratio A_1'/A_1 is called the amplitude coefficient of reflection and will be denoted by r. The ratio A_2/A_1 is called the amplitude coefficient of transmission and will be denoted by s.

The law of reflection,

$$\theta_1 = \theta_1',$$

implies that

$$\mu_1 = \cos\theta_1 = \cos\theta_1' = -\mu_1'.$$

Equations (7) and (8) can therefore be written

$$r = \frac{c_1\cos\theta_1 - c_2\cos\theta_2}{c_1\cos\theta_1 + c_2\cos\theta_2}, \tag{7′}$$

and

$$s = \frac{2\,c_1\cos\theta_1}{c_1\cos\theta_1 + c_2\cos\theta_2}. \tag{8′}$$

The case of normal incidence has been treated by Lamb (1932: 262).

Snell's law,

$$\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_2}{c_2},$$

makes it possible to obtain r as a function of c_1/c_2 for any given values of the angle of incidence θ_1. This relation is shown graphically in Fig. 8.

An important feature of the relation which can be seen from the graph is that r changes very little between an angle of incidence of 0° and 20°, when $c_1/c_2 > 1$, that is, when the incident waves approach a decrease in depth. Another noteworthy feature of the relation is that there is a critical angle of incidence when $c_1/c_2 < 1$, that is, when the incident wave approaches an increase in depth. If the angle of incidence exceeds this critical value, total reflection is indicated ($r = 1$).

The reflection coefficient has been derived on the assumption of an abrupt depth change. The theory of reflection of sound from a region between two parallel planes where the speed of propagation increases (or decreases) gradually indicates that the reflection coefficient depends on the ratio

$$\frac{d}{L/\sin\theta},$$

where d is the perpendicular distance between the planes, L the wave length, and θ the angle of incidence. If the ratio is small, the reflection coefficient is nearly as large numerically as it is for a discontinuous change. If the ratio is large, the reflection coefficient is nearly zero. Assuming the same criterion to apply to tsunamis, the reflection coefficient calculated above (Eq. 7′) is a fairly good approximation at steep slopes. For example, at the Continental Slope off northern California (near Crescent City) the depth increases from 3,000 to 9,900 ft. in about 15 mi.[4] (Fig. 1, inset). The wave length of a tsunami with 15 min. period in a depth of 9,900 ft. is 116 mi. For normal incidence the ratio

$$\frac{d}{L/\sin\theta} = \frac{15}{116} \doteq 0.13\,.$$

APPLICATION OF THEORY

Of various tide gauge records, that taken at Hanasaki can be used most satisfactorily to check the validity of the theoretically derived reflection coefficient. Since the direct waves were very small, the heights after the arrival of reflected waves are considered to result entirely from the reflected waves.

As noted above, the reflected waves which reached Hanasaki appear to have been reflected from the Continental Slope off southern Oregon. The height of waves reflected there can be estimated from the heights recorded at Cresent City (Fig. 9a), the nearest station from which

---

[4] Miles refers to statute miles throughout this paper.

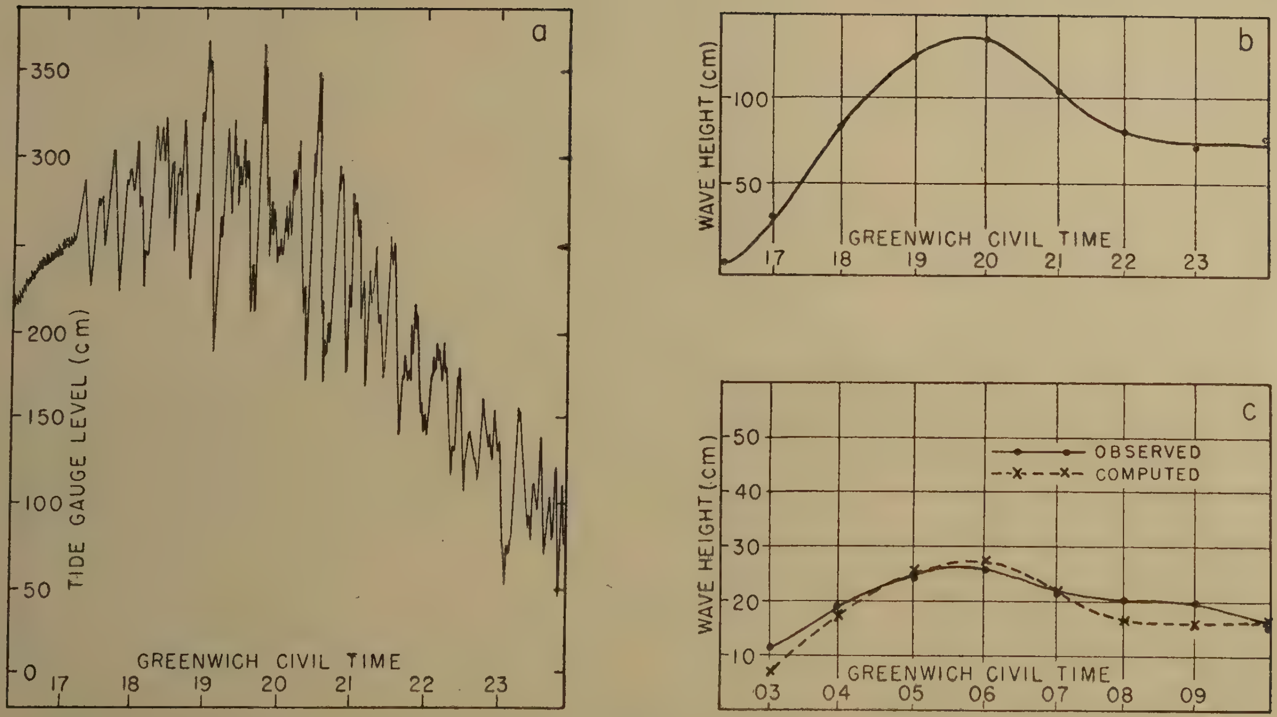


Figure 9. a. Crescent City tide gauge record for tsunami of April 1, 1946. b. Graph of average wave height for Crescent City. c. Comparison between observed and computed average wave heights at Hanasaki.

records are available. The heights averaged over two hour intervals, H_C, have been computed for each of the first 10 hours (Fig. 9b) in the same manner as they were for Hanasaki.

A typical bottom profile along a line seaward from near Crescent City is shown in the inset of Fig. 1. To facilitate computation, the bottom contours have been assumed to be straight lines parallel to the coastline. The actual contours do not differ widely from this picture.

Since the waves approach the coast almost at right angles, they diverge very little, and so Green's law (Lamb, 1932: 275),

$$\frac{H_1}{H_2} = \left(\frac{h_2}{h_1}\right)^{1/4},$$

can be used to obtain an estimate of the wave heights at 3,000 ft. from the recorded wave heights at Crescent City. The depth of the water at the tide well is assumed to be 20 ft. Thus the wave height at 3,000 ft., H_{3000}, is given by

$$H_{3000} = H_C\left(\frac{20}{3000}\right)^{1/4}. \tag{A}$$

The height of waves reflected from the slope at 9,900 ft. is obtained by applying the theory developed above. The reflection and transmission coefficients for incidence at 10° are nearly the same as those for normal incidence, namely

$$r = 0.27,$$

$$s = 1 + r = 1.27,$$

for $c_1/c_2 = \sqrt{9900/3000}$. The height of waves reflected at 9,900 ft., H_R, is then given by

$$H_R = \frac{r}{s} H_{3000} = \frac{0.27}{1.27} H_{3000}. \tag{B}$$

Reflection along the path (Fig. 1) of the reflected wave from the slope off southern Oregon to the Japanese Trench can be neglected, since there are no abrupt changes in depth.

The effect of distance D on the waves is obtained by using the relation

$$\frac{H_1}{H_2} = \left(\frac{D_2}{D_1}\right)^{1/3},$$

which applies to the first few waves of an impulsively generated wave train (Jeffreys, *et al.*, 1946: 485). In the present calculation D_1 is approximately 1,900 mi., and D_2 approximately 6,300 mi. The height off Hanasaki, at 9,900 ft., is given by

$$H_J = \left(\frac{1900}{6300}\right)^{1/3} H_R . \tag{C}$$

The height at the Hanasaki tide well, H_H, is obtained from H_J by Green's law. The depth of water at the tide well is again assumed to be 20 ft. Thus

$$H_H = \left(\frac{9900}{30}\right)^{1/4} H_J . \tag{D}$$

The average wave height at Crescent City during the first two hours of the tsunami is 85 cm. The calculation of the corresponding reflected wave height at Hanasaki is given in the table below:

(A) Wave height at 3,000 ft.,

$$H_{3000} = 85\left(\frac{20}{3000}\right)^{1/4} = 25 \text{ cm.}$$

(B) Height of wave reflected at 9,900 ft.,

$$H_R = 25\left(\frac{0.27}{1.27}\right) = 5.3 \text{ cm.}$$

(C) Wave height at 9,900 ft. off Hanasaki,

$$H_J = 5.3\left(\frac{1900}{6300}\right)^{1/3} = 3.6 \text{ cm.}$$

(D) Wave height at 20 ft. at Hanasaki,

$$H_H = 3.6\left(\frac{9900}{20}\right)^{1/4} = 17 \text{ cm.}$$

The travel time of the reflected wave from the Oregon Coast to Hanasaki is about 10 hours. The heights calculated in the above manner compare favorably with the average of the heights recorded at Hanasaki between 9 and 11 hours later (Fig. 9c). This close agreement is probably in part fortuitous, since many factors such as convergence and resonance have not been taken into account. The order of magnitude of the theoretical reflection coefficient, upon which the calculated height is critically dependent, seems to lead to agreement with this single set of observations.

The ratio between the average heights during the first 10 hours at Crescent City and the corresponding heights at Hanasaki remains nearly constant and serves to identify the waves at the two stations.

REFLECTION FROM CONTINENTAL SLOPES

Shepard (in press) gives 5° 20′ for the average slope and 10–20 mi. for the average range in width of the continental slopes bordering the Pacific Ocean. The depth at which the slope begins may vary from 600 to 3,000 ft. The present theory implies that the reflection coefficient at such average slopes is between 0.2 and 0.4 for normal incidence and between 0.15 and 0.3 for incidence at 45°. Thus, according to the theory, the reflected waves are of secondary but not negligible magnitude. At a given station, however, the reflected waves may be favored by convergence and be of primary magnitude. This might explain why the reflected waves are prominent only in relatively few cases.

REFLECTION FROM TRENCH WALLS

As pointed out above, the theory implies that a wave reaching a trench wall from shallower water is reflected totally if the angle of incidence exceeds a critical value. The arrival of the highest and most damaging waves at Napoopoo and Hokeena on the west coast of Hawaii about 18 hours after the April 1 earthquake (Shepard, *et al.*) can be explained by assuming that nearly total reflection occurred along the eastern edge of the Japan Trench and again at the Bonin Trench.

CONCLUSIONS

The data examined give good evidence that reflections do occur. One case has been subjected to quantitative treatment, and the observed wave heights were found to be in good agreement with those calculated on the basis of theory. The theory indicates that the reflected waves can be expected to be considerably smaller than the direct waves, except under very special circumstances. This conclusion is in agreement with observations.

REFERENCES

HART, G. E. F.
1931. Interesting oceanographic phenomena. Aust. Surveyor, December: 192–198.

JEFFREYS, HAROLD AND BERTHA S. JEFFREYS
1946. Methods of mathematical physics. Cambridge Univ. Press. 679 pp.

LAMB, HORACE
1932. Hydrodynamics. Cambridge Univ. Press, sixth ed. 738 pp.

SHEPARD, F. P.
Submarine geology. Harper and Bros., New York. 338 pp. (In press.)

SHEPARD, F. P., G. A. MACDONALD AND D. C. COX
The tsunami of April 1, 1946. Bull. Scripps Inst. Oceanogr., *5* (7). (In press.)

PROGRESS IN THE STATISTICAL THEORY OF TURBULENCE[1]

By

THEODORE VON KÁRMÁN

California Institute of Technology
Pasadena, California

The fundamental notion of statistical mean values in fluid mechanics was first introduced by Reynolds (1894). In his description of the turbulent motion he started from the molecular disorder. By considering the mean value and correlations in the molecular motion, as it was shown before, one obtains the equation of Stokes-Navier (1845, 1822). The solutions of these equations were designated by Reynolds as "mean motion." The second step in Reynolds' development is the consideration of the turbulent fluctuation. He assumes, for instance, that the apparently parallel flow in a cylindrical tube consists of a fluctuating motion characterized by the fact that the motion, at every instant, satisfies the Stokes-Navier equation. The parallel motion obtained by the formation of average values is called the "mean mean motion."

Reynolds' most important contributions were the definition of the mean values for the so-called Reynolds stresses and the recognition of the analogy between momentum transfer, transfer of heat and matter in the turbulent motion.

In the decades following Reynolds' discoveries, turbulence theory was directed to find semiempirical laws for the mean mean motion by methods borrowed from the kinetic theory of gases, *i. e.*, from the theory of the mean motion.

Prandtl's (1925, 1927) ideas on momentum transfer and Taylor's (1915, 1932) suggestions concerning vorticity transfer belong to the most important contributions of this period. But I believe that my formulation of the problem (1930) by the application of the similarity principle has the merit of being more general and independent of the methods of the kinetic theory of gases. This theory of mine led to the discovery of the logarithmic law of velocity distribution for the case of homologous turbulence. I call the turbulence homologous if the distribution of turbulence fluctuations and their correlations are identical at every point of the field. They differ only in scale.

---

[1] Presented at the Heat Transfer and Fluid Mechanics Institute, Los Angeles, California, June 23, 1948.

The next important step was the definition of isotropic turbulence by Taylor (1935, 1936). Apparently the case of homologous turbulence, *i. e.*, the shear motion, is too complicated for a fundamental attack by statistical methods. It was Taylor's fortunate idea to simplify the problem by the consideration of a uniform and isotropic field of turbulent fluctuation. Since such a field can be realized at least approximately in the wind tunnel, the possibility was given for an experimental check of the statistical ideas.

The next period in the development of the theory of turbulence was devoted to the analysis of the quantities which are accessible to measurements in a wind tunnel stream. These quantities are the correlation functions and the spectral function. The general mathematical analysis of the correlations was executed by Howarth and myself (von Kármán and Howarth, 1938). One has to consider five scalar functions $f(r)$, $g(r)$, $h(r)$, $k(r)$, $l(r)$. These functions determine all double and triple correlations between arbitrary velocity components observed at two points because of the tensorial character of the correlations. The two scalar functions for the double correlations are defined as follows:

$$f(r) = \frac{\overline{u_1(x_1, x_2, x_3)\, u_1\,(x_1 + r, x_2, x_3)}}{\overline{u_1^2}},$$

$$g(r) = \frac{\overline{u_1\,(x_1, x_2, x_3)\, u_1\,(x_1, x_2 + r, x_3)}}{\overline{u_1^2}}. \qquad (1)$$

Because of the continuity equation for incompressible fluids, $g = f + \frac{r}{2}\frac{df}{dr}$. For the same reason the triple correlations h, k, and l can be expressed by one of them, *i. e.*, by

$$h(r) = \frac{\overline{[u_1\,(x_1, x_2, x_3)]^2\, u_1\,(x_1 + r, x_2, x_3)}}{[\overline{u_1^2}]^{3/2}}. \qquad (1a)$$

In addition, we also deduced a differential equation from the Stokes-Navier equation which gives the relation between the time derivative of the function f and the triple correlation function h.

$$\frac{\partial}{\partial t}\,(\overline{u^2}f) + 2[\overline{u^2}]^{3/2}\left(\frac{\partial h}{\partial r} + \frac{4}{r}\,h\right) = 2\nu\overline{u^2}\left(\frac{\partial^2 f}{\partial r^2} + \frac{4}{r}\,\frac{\partial f}{\partial r}\right) \qquad (2)$$

We discussed this equation in two special cases:

a) Small Reynolds Number—in this case the triple correlations can be neglected and one obtains a self-preserving form for the double correlation function as function of r/λ, where λ is defined by the relation

$$\frac{d\overline{u^2}}{dt} = -\, 10\nu \frac{\overline{u^2}}{\lambda^2}\,. \tag{3}$$

b) Large Reynolds Number—in this case the terms containing the viscosity can be neglected for not too small values of r, and the functions f and h are assumed to be functions of the variable r/L, L being a length characterizing the scale of turbulence. The hypothesis of self-preserving correlation function leads to the following special results. One can consider three simple cases:

1. L = constant; then we have $\overline{u^2} \sim t^{-2}$ (Taylor).
2. Loitsianskii (1939) has shown that if the integral $\overline{u^2} \int_0^\infty f(r)\, r^4 dr$ exists, it must be independent of time and consequently $\overline{u^2}L^5$ = constant. Then $\overline{u^2} \sim t^{-10/7}$, $L \sim t^{2/7}$.
3. If the self-preserving character is extended to all values of r, *i. e.*, also near $r = 0$, one has $\overline{u^2} \sim t^{-1}$, $L \sim t^{1/2}$ (Dryden, 1943).

On the other hand, Taylor (1938) introduced a spectral function for the energy passing through a fixed cross section of a turbulent stream as the Fourier transform $\mathfrak{F}_0(n)$ of the correlation function $f(r)$. The relations between $\mathfrak{F}$ and f are given by the following equations:

$$\begin{aligned} f(r) &= \int_0^\infty \mathfrak{F}_0(n) \cos \frac{2\pi nr}{U}\, dn\,, \\ \mathfrak{F}_0(n) &= \frac{4\overline{u_1{}^2}}{U} \int_0^\infty f(r) \cos \frac{2\pi nr}{U}\, dr\,. \end{aligned} \tag{4}$$

In these equations n is the frequency of the fluctuation of the uniform velocity U as function of time. Relative to the stream, $\mathfrak{F}_0(n)$ can be replaced by $\mathfrak{F}_1(\kappa_1)$, where $\kappa_1 = \dfrac{2\pi n}{U}$, *i. e.*, the wave number of the fluctuation, measured in the x_1-direction.

It is seen that in this period of the development of the turbulence theory the analytical and experimental means for the study of isotropic turbulence were clearly defined, but with the exception of the case of very small Reynolds numbers no serious attempt was made to find the laws for the shapes of either the correlation or the spectral functions. I believe this is the principal aim of the period in which we find ourselves at present. Promising beginnings were made by Kolmogoroff (1941), Onsager (1945), Weizsäcker (1946), and Heisenberg (1946). I do not want to follow the special arguments of these authors. Rather, I want to define the problem clearly and point out the relations between assumptions and results.

First, we will assume that the three components of the velocity in an homogeneous isotropic turbulent field, at any instant, can be developed in the manner of Fourier's integrals

$$u_i = \iiint_{-\infty}^{\infty} Z_i(\kappa_1, \kappa_2, \kappa_2, t) e^{i(\kappa_1 x_1 + \kappa_2 x_2 + \kappa_3 x_3)} \, d\kappa_1 d\kappa_2 d\kappa_3 \,. \tag{5}$$

Second, the intensity of the turbulent field shall be characterized by the quadratic mean value $\overline{u_i^2}$ (level of the turbulence). Also, there exists a function $\mathfrak{F}(\kappa)$, such that $[\overline{u_i^2}]_0^\kappa = \int_0^\kappa \mathfrak{F}(\kappa')\, d\kappa'$, where the symbol $[\overline{u_i^2}]_0^\kappa$ means a partial mean value of the square of the velocity, the averaging process being restricted for such harmonic components whose wave numbers κ_1, κ_2, κ_3 satisfy the relation

$$\kappa_1^2 + \kappa_2^2 + \kappa_3^2 \leqq \kappa^2 \,.$$

If such a function exists, it is connected with the spectral function of Taylor $\mathfrak{F}_1(\kappa_1)$ by the relation

$$\mathfrak{F}_1(\kappa_1) = \tfrac{1}{4} \int_{\kappa_1}^{\infty} \frac{1}{\kappa^3} \mathfrak{F}(\kappa)\, (\kappa^2 - \kappa_1^2)\, d\kappa \,. \tag{6}$$

This relation was found by Heisenberg (1946). It expresses the geometrical fact that all oblique waves (Fig. 1) with wave length

$$\frac{2\pi}{\kappa} < \frac{2\pi}{\kappa_1}$$

necessarily contribute in the one-dimensional analysis to the waves with wave length

$$\frac{2\pi}{\kappa_1} \,.$$

Third, it is evident that there must be an equation for the time derivative of $\mathfrak{F}(\kappa)$ which corresponds to the equation for the time derivative of $f(r)$ which has been found by Howarth and myself.

The physical meaning of this equation is evident. Let us start from the energy equation

$$\frac{1}{2}\frac{\partial \overline{u_i^2}}{\partial t} + \overline{\left(u_i u_j + \delta_{ij}\frac{p}{\rho}\right)\frac{\partial u_i}{\partial x_j}} = \nu \frac{\partial^2 u_i}{\partial x_j^2} u_j \, . \tag{7}$$

The right side represents the energy dissipation by viscous forces. The second term on the left side is the work of the Reynolds stresses; it represents a transfer of energy without actual dissipation. Our problem is to find $\frac{\partial \mathfrak{F}}{\partial t}$ by Fourier analysis and averaging process. One finds the contribution of the viscous forces to be equal to $-2\nu\,\mathfrak{F}(\kappa)\kappa^2$. Hence we write formally

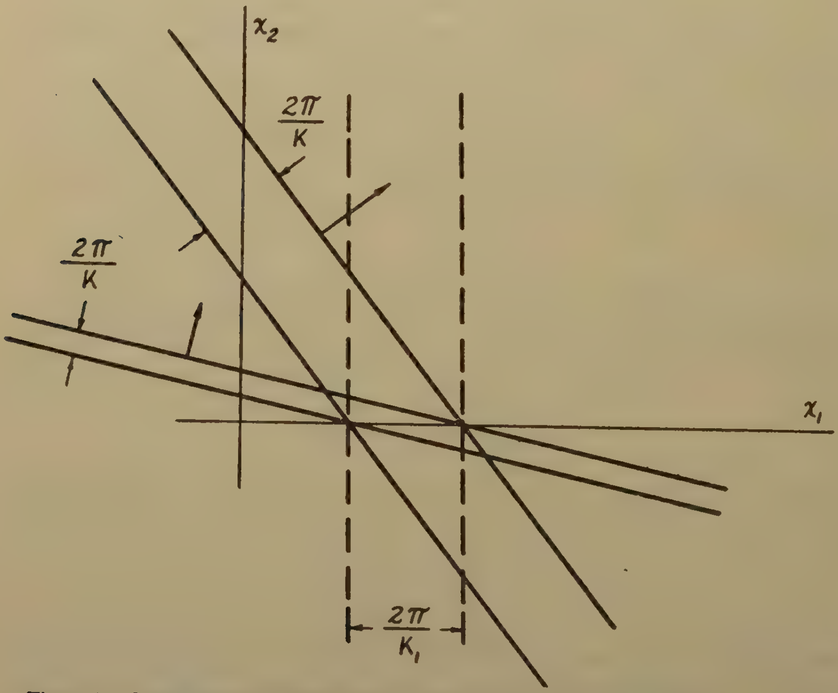


Figure 1. Contribution of oblique waves to plane waves in direction of x_1.

$$\frac{\partial \mathfrak{F}}{\partial t} + \mathcal{W}_\kappa = -2\,\nu\kappa^2\,\mathfrak{F}(\kappa)\, . \tag{8}$$

Here $\mathcal{W}_\kappa d\kappa$ is the balance for the energy contained in harmonic components contained in the interval $d\kappa$; obviously $\int_0^\infty \mathcal{W}_\kappa d\kappa = 0$. C. C. Lin (personal communication) has shown that $\mathcal{W}_\kappa = 2\mathcal{H}\kappa^2$, where $\mathcal{H}(\kappa) = 2\left(\kappa^2 \mathcal{H}_1''(\kappa) - \kappa \mathcal{H}_1'(\kappa)\right)$ and

$$\mathcal{H}_1(\kappa) = \frac{2\overline{u^2}^{3/2}}{\pi} \int_0^\infty h(r) \frac{\sin(\kappa r)}{\kappa} dr .$$

Unfortunately this relation does not help, as far as the determination of f and h is concerned. For example, if one expresses h in terms of f by means of the Kármán-Howarth equation, calculates $\mathcal{W}_\kappa$ and then substitutes the result in equation (8), one obtains an identity. It appears that at the present time one needs some additional physical assumption.

Fourth, we assume that $\mathcal{W}_\kappa$ can be expressed in the form;

$$\mathcal{W}_\kappa = \int_0^\infty \Theta\{\mathfrak{F}(\kappa), \mathfrak{F}(\kappa'), \kappa, \kappa'\}\, d\kappa' . \tag{9}$$

The physical meaning of this assumption is the existence of a transition function for energy between the intervals $d\kappa$ and $d\kappa'$ which depends only on the energy density and the wave numbers of the two intervals. It follows from this definition that by interchanging κ and κ' one has

$$\Theta\{\mathfrak{F}(\kappa), \mathfrak{F}(\kappa'), \kappa, \kappa'\} = -\Theta\{\mathfrak{F}(\kappa'), \mathfrak{F}(\kappa), \kappa', \kappa\} . \tag{10}$$

It must be noted that our assumption probably cannot be exact. It is very probable that the values of $\mathfrak{F}$ for the difference and the sum of κ and κ' also enter in the transition function. I believe that the assumption gives a fair approximation when κ and κ' are very different, but it is certainly untrue if κ and κ' are nearly equal.

Fifth, we furthermore specify the function Θ in the following way;

$$\Theta = -C\, \mathfrak{F}^{\alpha}(\kappa)\, \mathfrak{F}^{\alpha'}(\kappa')\, \kappa^{\beta} \kappa'^{\beta'} ; \quad C = \text{const.} \tag{11}$$

It follows from dimensional reasoning that

$$\alpha + \alpha' = 3/2 , \qquad \beta + \beta' = 1/2 .$$

As a result of the sequence of assumptions given above, we obtain the equation;

$$\frac{\partial \mathfrak{F}}{\partial t} = C\left[\mathfrak{F}^{\alpha} \kappa^{\beta} \int_0^\kappa \mathfrak{F}^{3/2-\alpha}(\kappa') \kappa_1{}^{1/2-\beta} d\kappa' - \mathfrak{F}^{3/2-\alpha} \kappa^{1/2-\beta} \int_\kappa^\infty \mathfrak{F}^{\alpha}(\kappa') \kappa'^{\beta} d\kappa' \right] - 2\nu\kappa^2 \mathfrak{F} . \tag{12}$$

Obviously, if $\mathfrak{F}$ is known for $t = 0$, equation (12) determines the values of $\mathfrak{F}$ for all times. If one neglects the first term on the left side, which represents the decay of turbulence, and chooses the specific values $\alpha = 1/2$, $\beta = -3/2$, one arrives at the theory proposed by Heisenberg (1946).

Let us consider the case of large Reynolds numbers but assume that κ is not so large that the term containing the viscosity coefficient becomes significant. Let us assume also that the first term on the left side is small by comparison to the second term. Physically, this means that the energy entering in the interval $d\kappa$ is equal to the energy which leaves the interval. Then one has the relation:

$$\mathfrak{F}^{\alpha}\kappa^{\beta}\int_0^{\kappa}\mathfrak{F}^{3/4-\alpha}(\kappa')\kappa'^{1/2-\beta}d\kappa' = \mathfrak{F}^{3/2-\alpha}\kappa^{1/2-\beta}\int_{\kappa}^{\infty}\mathfrak{F}^{\alpha}(\kappa')\kappa'^{\beta}d\kappa' \,. \tag{13}$$

This equation is satisfied by the solution $\mathfrak{F}(\kappa) \sim \kappa^{-5/3}$, as one can easily see by substitution in equation (13). This result is independent, evidently, of the special choice of α and β. That is the reason why it was independently found by Onsager (1945), Kolmogoroff (1941) and Weizsäcker (1946). It is essentially a consequence of dimensional considerations. Let us now stay with the case of large Reynolds numbers by neglecting again the viscosity term while retaining the first term on the left side. In other words, we consider the actual process of decay at large Reynolds numbers. Let us assume that $\mathfrak{F}$ is a function of a nondimensional variable κ/κ_0, when κ_0 is a function of time. This assumption is equivalent to our former assumption that $f(r)$ is a function of r/L; *i. e.*, we assume that $\mathfrak{F}$ and f preserve their shapes during the decay. Evidently $\kappa_0 \sim 1/L$. Then the function can be written in the form

$$\mathfrak{F}(\kappa) = \frac{\overline{u^2}}{\kappa_0}\,\Phi\left(\frac{\kappa}{\kappa_0}\right).$$

Then with $\kappa/\kappa_0 = \xi$ and

$$\frac{\partial\mathfrak{F}}{\partial t} = \frac{\Phi}{\kappa_0}\frac{d\overline{u^2}}{dt} - \frac{\overline{u^2}}{\kappa_0^{\,2}}\,\Phi\,\frac{d\kappa_0}{dt} - \frac{\overline{u^2}}{\kappa_0^{\,2}}\,\Phi'\,(\xi)\,\xi\,\frac{d\kappa_0}{dt}\,,$$

equation (12) becomes

$$\left(\frac{1}{\kappa_0}\frac{d\overline{u^2}}{dt} - \frac{\overline{u^2}}{\kappa_0^{\,2}}\frac{d\kappa_0}{dt}\right)\Phi - \frac{\overline{u^2}}{\kappa_0^{\,2}}\frac{d\kappa_0}{dt}\,\Phi'\,\xi + \mathfrak{W}_{\kappa} = 0\,, \tag{14}$$

where

$$\mathfrak{W}_{\kappa} = -\,C\,[\overline{u^2}]^{3/2}\,[\Phi^{\alpha}\xi^{\beta}\int_0^{\xi}\Phi(\xi')^{3/2-\alpha}\xi'^{1/2-\beta}d\xi' \\ - \Phi^{3/2-\alpha}\xi^{1/2-\beta}\int_{\xi}^{\infty}\Phi(\xi')^{\alpha}\,\xi'^{\beta}d\xi']\,.$$

According to Loitsianskii's (1939) results,

$$\frac{1}{\overline{u^2}}\frac{d\overline{u^2}}{dt} = \frac{5}{\kappa_0}\frac{d\kappa_0}{dt},$$

and one obtains the equation

$$\xi^5\,(\xi^{-4}\Phi)' = -\,5C\,\frac{[\overline{u^2}]^{3/2}}{\dfrac{d\overline{u^2}}{dt}}\left[\Phi^{\alpha}\xi^{\beta}I_0{}^{\xi} - \Phi^{3/2-\alpha}\xi^{1/2-\beta}I_{\xi}{}^{\infty}\right], \tag{15}$$

where

$$I_0{}^{\xi} = \int_0^{\xi}\Phi(\xi')^{3/2-\alpha}\xi'^{1/2-\beta}d\xi'\,;\quad I_{\xi}{}^{\infty} = \int_{\xi}^{\infty}\Phi(\xi')^{\alpha}\xi'^{\beta}d\xi'\,.$$

Let us assume that $4\alpha + \beta < 5/2$ as, for example, in the case of Heisenberg (1946). Then for small values of ξ the right side of equation (15) is small in comparison with the term on the left side, and one has

$$\Phi\,(\xi) \cong \text{const}\;\xi^4\,.$$

If $4\alpha + \beta > 5/2$, $\mathfrak{F}$ begins with a lower power of κ than κ^4, one can show that the integral $\int_0^{\infty} r^4 f(r)dr$ does not converge, so that Loitsianskii's result is incorrect. I should like to investigate this second case in a later work. Let us assume, for the time being, that Loitsianskii's result is correct; therefore the first case prevails. Then it follows that $\mathfrak{F}$ or Φ behaves as $(\kappa/\kappa_0)^4$ for small values of κ and is proportional to $(\kappa/\kappa_0)^{-5/3}$ for large values of κ. For any definite choice of α and β the differential equation (15) can be solved numerically. The result that $\mathfrak{F} \cong \kappa^4$ for small values of κ was also found in a different way by C. C. Lin (personal communication).

For the time being I propose an interpolation formula as follows:

$$\Phi\,(\xi) = \text{const.}\,\frac{\xi^4}{(1+\xi^2)^{17/6}}\,. \tag{16}$$

This interpolation formula represents correctly $\Phi\,(\xi)$ for small and large values of ξ and has the advantage that all calculations can be carried out analytically by use of known functions. The results are as follows:

$$
\begin{aligned}
\mathfrak{F}(\kappa/\kappa_0) &= \text{const.}\,\frac{(\kappa/\kappa_0{}^4)}{[1 + (\kappa/\kappa_0)^2]^{17/6}}\,, \\
\mathfrak{F}_1(\kappa_1/\kappa_0) &= \text{const.}\,\frac{1}{[1 + (\kappa_1/\kappa_0)^2]^{5/6}}\,, \\
f(\kappa_0 r) &= \frac{2^{2/3}}{\Gamma(1/3)}\,(\kappa_0 r)^{1/3}\,K_{1/3}\,(\kappa_0 r)\,, \\
g(\kappa_0 r) &= \frac{2^{2/3}}{\Gamma(1/3)}\,(\kappa_0 r)^{1/3}\left[K_{1/3}\,(\kappa_0 r) - \frac{\kappa_0 r}{2}\,K_{-2/3}\,(\kappa_0 r)\right].
\end{aligned}
\tag{17}
$$

The K's are Bessel functions with imaginary argument. For small values of $\kappa_0 r$,

$$
f(\kappa_0 r) = 1 - \frac{\Gamma(2/3)}{\Gamma(4/3)}\left(\frac{\kappa_0 r}{2}\right)^{2/3}, \tag{18}
$$

as suggested by Kolmogoroff's theory.

I have compared these results with the measurements of Liepmann, Laufer and Liepmann (1948) carried out at the California Institute of Technology with the financial assistance of the National Advisory Committee for Aeronautics.[2] These observations were made in the 10-foot wind tunnel of the Guggenheim Aeronautical Laboratory, using a grid whose mesh size was $M = 4''$. The measurements were made at a distance $x = 40.4M$ from the grid. Fig. 2 shows the comparison of calculated and measured values for the spectral function $\mathfrak{F}_1(\kappa_1)$. It has to be taken into account that the observed values of $\mathfrak{F}_1(\kappa_1)$ have large scatter; the deviation for high values of κ_1 corresponds to the beginning influence of viscosity. Fig. 3 gives a comparison between measured and calculated values of the correlation function $g(r)$. This function is chosen because the observations are more accurate than in any other case. It is seen that the agreement is almost too good in view of the assumptions made above. One must remark that there is only one arbitrary constant in the formula for g, *viz.*, the constant κ_0 which determines the scale of the turbulence. It is true that some of the data of Liepmann, *et al.* (1948) do not show such a good agreement. The agreement is excellent for values of g larger than 0.1, but after that the measured values are higher than the calculated ones. Possibly some oscillations existing in the wind tunnel stream were interpreted as turbulence, or the turbulence is not quite isotropic.

---

[2] The N. A. C. A. has kindly granted permission for these data to be used here before appearing in its official publications.

I believe that the merits of my deduction are:

a) The assumptions involved are exactly formulated;
b) The specific assumptions of Heisenberg's theory concerning the transition function are not used;
c) The actual process of decay is considered;
d) The analysis is extended to the lower end of the turbulence spectrum.

Concerning the case of large values of κ (small values of r), Kovászнay (1948) introduced an interesting assumption which is more

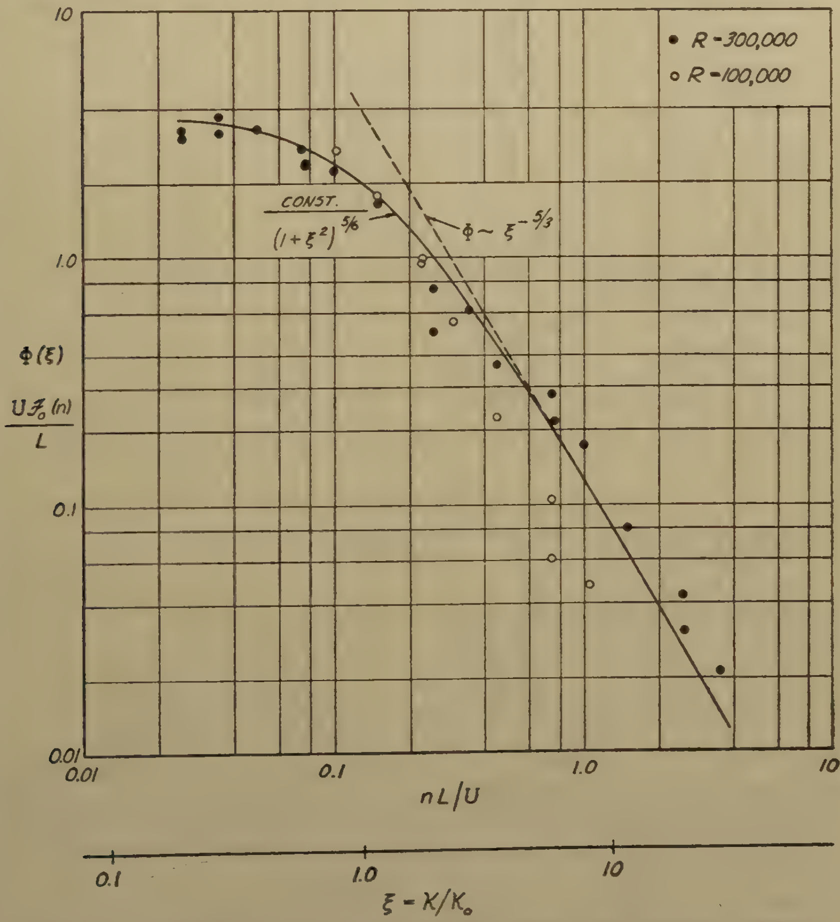


Figure 2. Comparison of observed and computed values of the frequency spectrum.

restricting than my fourth assumption. Obviously $\int_0^\kappa \mathcal{W}_\kappa d\kappa$ is the total energy transferred by the Reynolds stresses from the interval $0 \rightarrow \kappa$ to the interval $\kappa \rightarrow \infty$. Kovásznay assumes—following Kolmogoroff's (1941) arguments—that this quantity is a function of $\mathfrak{F}(\kappa)$ and κ only. Then for dimensional reasons

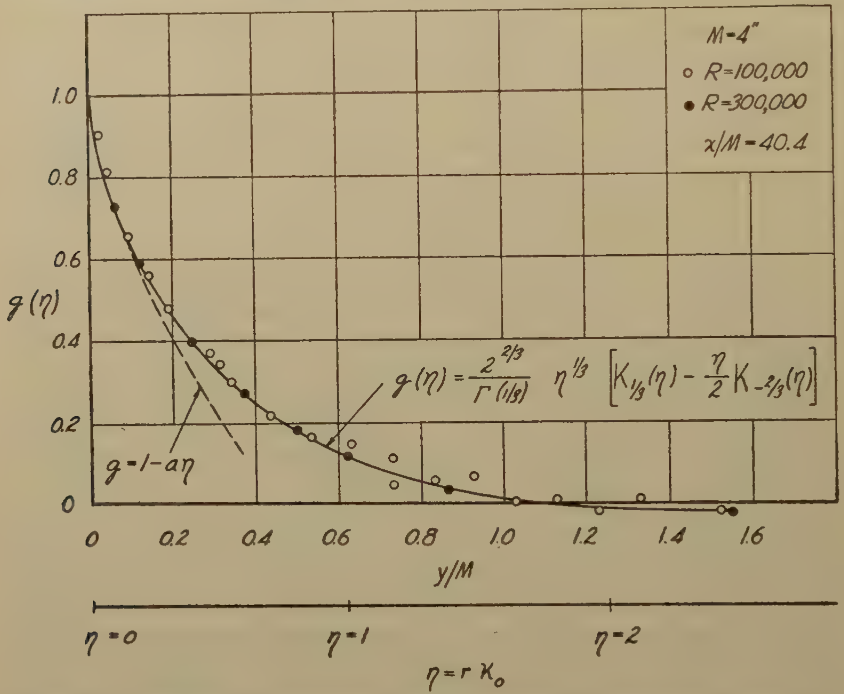


Figure 3. Comparison of observed and computed values of the correlation function $g(r)$. The Reynolds number is based on the stream velocity and the mesh size.

$$\int_0^\kappa \mathcal{W}_\kappa d\kappa = \text{const } \mathfrak{F}^{3/2}\kappa^{5/2}.$$

This assumption appears to be correct for large values of κ. However, when the assumption is extended to the range of small values of κ, and one substitutes $\mathcal{W}_\kappa$ in equation (8), one can calculate easily $\mathfrak{F}(\kappa)$. Neglecting the viscous term, one obtains the relation

$$\mathfrak{F}(\kappa/\kappa_0) = \text{const.} \frac{(\kappa/\kappa_0)^4}{[1 + \mathfrak{F}^{1/2}(\kappa/\kappa_0)^{3/2}]^{17/2}}. \tag{19}$$

The right side of equation (19) behaves as my corresponding equation (17) for small and large values of κ/κ_0. It will be interesting to see how far different transition from small to large values influences the accordance with observation.

REFERENCES

Dryden, H. L.

1943. A review of the statistical theory of turbulence. Quart. appl. Math., *1:* 7-42.

Heisenberg, W.

1946. Zur Statistischen Theorie der Turbulenz. Unpublished.

Kármán, Theodore von

1930. Mechanische Ähnlichkeit und Turbulenz. Nachr. Ges. Wiss. Gottingen, Mat.-phys. Kl., pp. 58–76.

Kármán, Theodore von, and Leslie Howarth

1938. On the statistical theory of isotropic turbulence. Proc. roy. Soc. Lond., (A) *164:* 192–215.

Kovásznay, L. S. G.

1948. The spectrum of locally isotropic turbulence. Phys. Rev., *73* (9): 1115.

Kolmogoroff, A.

1941. The local structure of turbulence in incompressible viscous fluids for very large Reynolds numbers. C. R. (Doklady) Acad. Sci. U. R. S. S., N. S. *30* (4): 301–305.

Liepmann, H. W., J. Laufer and K. Liepmann

1948. On some turbulence measurements behind grids. Final Rep. nat. adv. Comm. Aero., Wash. Contract NAw 5442.

Loitsianskii, L. G.

1939. Some basic laws of isotropic turbulent flow. Rep. Cent. Aero-Hydrodynam. Inst., Moscow, No. 440. (Transl., Techn. Memo. nat. adv. Comm. Aero., Wash., 1079: 1–36, 1945.)

Navier, M. H.

1822. Sur les lois du mouvement des fluides. Mem. Acad. Sci., Paris, *6:* 389–440.

Onsager, Lars

1945. The distribution of energy in turbulence. Phys. Rev., (II) *68:* 286.

Prandtl, L.

1925. Bericht über Untersuchungen zur ausgebildeten Turbulenz. Z. angew. Math. Mech., *5:* 136–139.

1927. Ueber die ausgebildete Turbulenz. Verh. 2. int. Kongr. techn. Mechanik, Zurich, 1926, pp. 62–74.

Reynolds, Osborne

1894. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. Proc. roy. Soc. Lond., *56:* 40–45 (Abstr.).

Stokes, G. G.

1845. On the theories of the internal friction of fluids in motion, and of the equilibrium and motion of elastic solids. Trans. Camb. phil. Soc., *8:* 287–319.

TAYLOR, G. I.

1915. Eddy motion in the atmosphere. Philos. Trans., (A) 215: 1–26.

1932. The transport of vorticity and heat through fluids in turbulent motion. Proc. roy. Soc. Lond., (A) *135:* 685–702.

1935. Statistical theory of turbulence. I. Proc. roy. Soc. Lond., (A) *151:* 421–444.

1935. Statistical theory of turbulence. II. Proc. roy. Soc. Lond., (A) *151:* 444–454.

1935. Statistical theory of turbulence. III. Distribution of dissipation of energy in a pipe over its cross-section. Proc. roy. Soc. Lond., (A) *151:* 455–464.

1935. Statistical theory of turbulence. IV. Diffusion in a turbulent air stream. Proc. roy. Soc. Lond., (A) *151:* 465–478.

1936. Statistical theory of turbulence. V—Effect of turbulence on boundary layer. Theoretical discussion of relationship between scale of turbulence and critical resistence of spheres. Proc. roy. Soc. Lond., (A) *156:* 307–317.

1938. The spectrum of turbulence. Proc. roy. Soc. Lond., (A) *164:* 476–490.

WEIZSÄCKER, C. F. VON

1946. Das Spektrum der Turbulenz bei grossen Reynolds' schen Zahlen. Unpublished.

AN ANALYSIS OF THE STIRRING AND MIXING PROCESSES IN INCOMPRESSIBLE FLUIDS

BY

CARL ECKART

Scripps Institution of Oceanography[1]
University of California
La Jolla, California

INTRODUCTION

In the laboratory, the phenomena accompanying the mixture of two fluids are so commonplace that they do not seem to merit investigation. In the ocean and atmosphere, the phenomena are presumably very similar, but the lack of a good theory on which to base calculations complicates the problem much more than in the laboratory.

It is useful to consider a trivial experiment by way of introduction: the mixing of coffee and cream. Three more or less distinct stages can be observed:

1. *The initial stage,* in which rather large volumes of cream and coffee are distinctly visible; there are sharp gradients at the interfaces between the volumes, but elsewhere the gradient is practically zero. Averaged over the entire volume, the gradient is small. If motion of the liquids is avoided, this state persists for a considerable time.
2. *The intermediate stage,* after motion has been induced by stirring the liquids; the masses of cream and coffee are distorted, with a rapid increase in the extent of the interfacial regions having high concentration gradients. The average value of the gradient is correspondingly increased.
3. *The final stage,* in which the gradients disappear, apparently quite suddenly and spontaneously, with the liquid becoming homogeneous.

It is a reasonable working hypothesis to assume that these three stages (or at least the second and third) also occur in the ocean and atmosphere when concentration or temperature differences arise.

The potentially long duration of the first stage may be ascribed to the slowness of the diffusion or heat conduction process which results from the small average value of the gradients. During the second stage, the average value of the gradient is increased by the relative motion of different parts of the liquid as a result of stirring. The disappearance of the gradients during the third stage is presumably caused by molecular diffusion. This is more rapid than during the first stage, because stirring has increased the gradients and thus has also increased the diffusion currents. For convenience, this last stage may be called mixing, to distinguish it from stirring.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 389.

TWO EXAMPLES OF STIRRING

As a first example of stirring, one may consider the formation of a steep thermocline by lateral advection; Fig. 1 represents a vertical section of the fluid. Suppose that initially there is a small horizontal gradient, but that the regions *A* and *B* (Fig. 1a) are at appreciably different temperatures because of the relatively large distance separating them. Suppose further that the fluid everywhere moves horizontally, but that there is a vertical velocity gradient; the motion is laminar, with shear across the horizontal planes. Then, since the upper parts of the regions *A* and *B* move with a higher velocity than the lower parts, these regions will be distorted until eventually the configuration of Fig. 1b is produced. The particles 1 and 2 will have moved much closer together, although the temperature difference between them is unaltered. Thus the original small horizontal gradient will have been converted into a much larger vertical gradient.

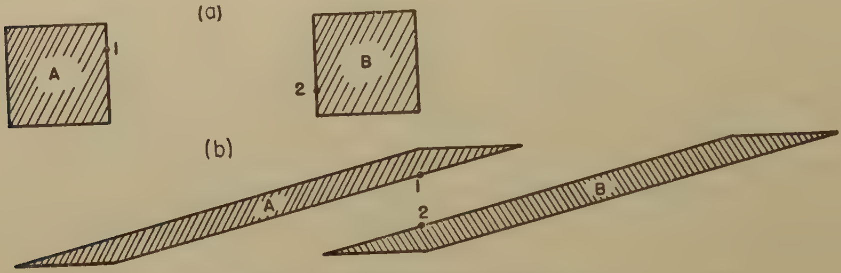


Figure 1. Stirring action of laminar flow.

This example suggests that the large thermal gradients shown on the bathythermogram of Fig. 2 may be indicative of regions of relatively large velocity gradients, and that the whole trace is indicative of horizontal shear in the layers between the depth of 120 and 240 feet. Possibly this suggestion could be tested experimentally.

As a second example, suppose the initial temperature distribution to be as before, but that an eddy is somehow set up with its axis horizontal and perpendicularly bisecting the line joining the centers of *A* and *B*

(Fig. 3a). In this eddy, let the orbits of the fluid particles be circular and be described with an angular velocity ω that is a function of their distance from the axis of the eddy. Then, after the lapse of successive intervals of time, the regions A and B of Fig. 3a will be deformed and assume the configuration shown in Figs. 3b and 3c. It is clear that such an eddying motion will be very effective, not merely in increasing the mean gradient but also in increasing the interfacial area of the regions A and B and in decreasing the interfacial distance between them. All three effects are such as to increase the flow of heat by conduction. During the stage of stirring these three processes dominate; ultimately they may so increase the conduction of heat that the latter dominates and causes the mean gradient to diminish toward zero.

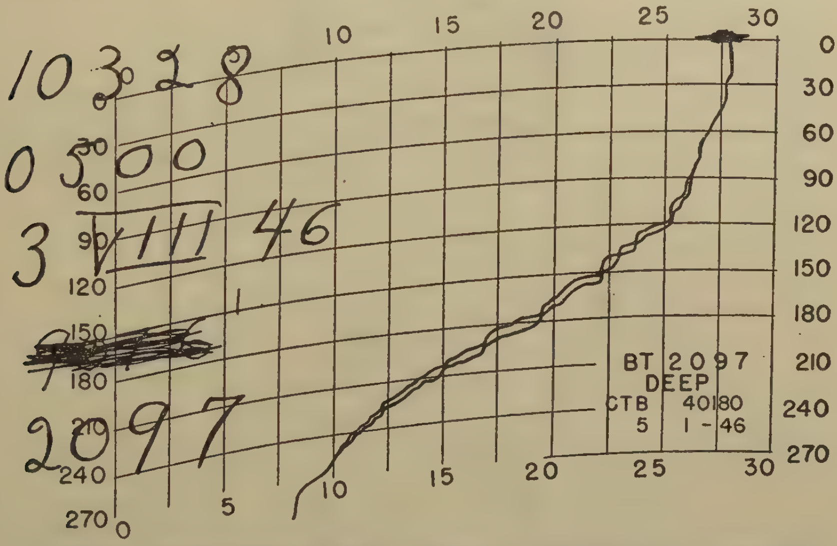


Figure 2. Bathythermogram taken on August 3, 1946, from U. S. S. Ingraham at Lat. 11° 48′ N., Long. 164° 46′ E. Temperature in °F., depth in feet.

It is not difficult to show that, because of the stirring action, the mean gradient at a distance r from the center of the eddy is

$$\tfrac{1}{2}\,|\,ar\,(d\omega/dr)\,t\,|\,,$$

where a is the initial horizontal gradient, and t is the elapsed time since the initial instant. Thus, if the eddy persists long enough, the gradient, and consequently the heat flow, may become very large indeed.

Having considered these two examples of stirring, the next section will examine the laws according to which heat conduction or diffusion operates to reduce the mean gradient. In the succeeding section, the manner in which these laws are modified by stirring will be treated in some generality.

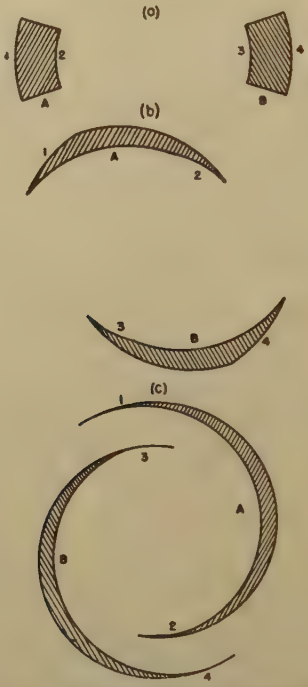


Figure 3. Stirring action of circular eddy.

THE EQUALIZATION OF GRADIENTS BY CONDUCTION OR DIFFUSION

In order to avoid, initially, the complication of the stirring process, consider conduction or diffusion in a solid substance. The differential equation governing this is

$$\frac{\partial \theta}{\partial t} = \kappa \nabla^2 \theta , \tag{1}$$

where κ is the conductivity or diffusion constant, and θ the temperature or concentration. On differentiating this, one obtains

$$\frac{\partial}{\partial t} \nabla \theta = \kappa \nabla^2 \nabla \theta ; \tag{2}$$

multiplying scalarly by $\nabla\theta$, the result is

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial t}(\nabla\theta)^2 &= \kappa\,\nabla\theta\cdot\{\nabla^2\,\nabla\theta\}\\ &= \kappa\,\nabla\cdot\{\nabla\theta\,\nabla^2\theta\} - \kappa\,\{\nabla^2\theta\}^2\\ &= \nabla\cdot\left\{\frac{\partial\theta}{\partial t}\nabla\theta\right\} - \kappa\left\{\nabla^2\theta\right\}^2.\end{aligned} \tag{3}$$

If one integrates the terms of this equation over a volume and uses the divergence theorem, the result is

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\iiint\limits_{vol}(\nabla\theta)^2\,d\tau &=\\ \iint\limits_{bndry}\nu\cdot\nabla\theta\frac{\partial\theta}{\partial t}\,d\sigma &- \kappa\iiint\limits_{vol}\{\nabla^2\theta\}^2\,d\tau\,,\end{aligned} \tag{4}$$

where $d\tau$ and $d\sigma$ are the elements of volume and area, and ν the unit vector normal to $d\sigma$.

At first, assume that the surface integral vanishes. This will certainly be the case if either: (a) the temperature of each point of the boundary is constant, so that $\partial\theta/\partial t = 0$, or (b) no heat flows through any point of the boundary, so that $\nu\cdot\nabla\theta = 0$. Then, since $\{\nabla^2\theta\}^2 \geqslant 0$, the right side of (4) is either negative or at most zero. Hence the integral on the left must decrease, or at most remain constant.

Define the rms gradient, G, by the equation

$$G^2 = \frac{1}{V}\iiint\limits_{vol}(\nabla\theta)^2\,d\tau\,, \tag{5}$$

where V is the volume of the region under consideration. The quantity I, defined by

$$I^2 = \frac{1}{V}\iiint\limits_{vol}(\nabla^2\theta)^2\,d\tau\,, \tag{6}$$

is a convenient measure of the inhomogeneity of the material inside V. Using these definitions,

$$\frac{1}{2}\frac{dG^2}{dt} = -\kappa I^2\,. \tag{7}$$

For a given inhomogeneity, therefore, the mean square gradient decreases at a rate *directly* proportional to the constant κ.

The activity of the material inside V is measured by the quantity A, defined by

$$A^2 = \frac{1}{V} \iiint\limits_{vol} \left(\frac{\partial \theta}{\partial t}\right)^2 d\tau \, . \tag{8}$$

Because of (1), (7) may also be written

$$\frac{1}{2} \frac{dG^2}{dt} = - \frac{1}{\kappa} A^2 \, . \tag{7.1}$$

For a given activity, the mean square gradient thus decreases at a rate *inversely* proportional to the constant κ.

However, the most important conclusion is that whenever the surface integral of (4) vanishes the rms value of the gradient never increases. In the case of a small solid object in the laboratory, the conditions for the surface integral to vanish are easily fulfilled. However, if an arbitrary volume of the ocean or the atmosphere is considered, this is no longer true. The conduction of heat through the boundary, and its consequent change in temperature, may become very important. Moreover, the process of stirring must be considered.

MODIFIED EQUALIZATION THEOREMS WHEN THE SUBSTANCE IS AN INCOMPRESSIBLE FLUID

When the substance under consideration is an incompressible fluid, rather than a solid, (1) becomes

$$\frac{D\theta}{Dt} = \kappa \nabla^2 \theta \, , \tag{10}$$

or, using the indicial notation,

$$\frac{D\theta}{Dt} = \kappa \frac{\partial^2 \theta}{\partial x_i^2} \, , \tag{11}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u_i \frac{\partial}{\partial x_i} \, .$$

The differentiation of this equation (as above) encounters a complication, because

$$\frac{\partial}{\partial x_j} \frac{D}{Dt} = \frac{D}{Dt} \frac{\partial}{\partial x_j} + \frac{\partial u_i}{\partial x_j} \frac{\partial}{\partial x_i} \, . \tag{12}$$

The analogue of (2) is therefore

$$\frac{D}{Dt}\frac{\partial\theta}{\partial x_j} = \kappa\frac{\partial}{\partial x_j}\frac{\partial^2\theta}{\partial x_i^2} - \frac{\partial u_i}{\partial x_j}\frac{\partial\theta}{\partial x_i}, \tag{13}$$

and involves the velocity gradient $\partial u_i/\partial x_j$.

Multiplying by $\partial\theta/\partial x_j$, and summing,

$$\frac{1}{2}\frac{D}{Dt}\left(\frac{\partial\theta}{\partial x_i}\right)^2 = \kappa\frac{\partial}{\partial x_j}\left(\frac{\partial\theta}{\partial x_j}\frac{\partial^2\theta}{\partial x_i^2}\right) - \kappa\left\{\frac{\partial^2\theta}{\partial x_i^2}\right\}^2 - \frac{\partial u_i}{\partial x_j}\frac{\partial\theta}{\partial x_i}\frac{\partial\theta}{\partial x_j}. \tag{14}$$

In order to obtain simple integrated results, it is now necessary to choose a volume V whose boundary is not stationary but which moves so that it always encloses the same fluid. Integrating over such a moving volume, the result is

$$\frac{1}{2}\frac{dG^2}{dt} = \iint_{bndry} \nu\cdot\nabla\theta\frac{D\theta}{Dt}\,d\sigma - \kappa I^2 - S, \tag{15}$$

where the additional term is

$$S = \iiint_{vol}\frac{\partial u_i}{\partial x_j}\frac{\partial\theta}{\partial x_i}\frac{\partial\theta}{\partial x_j}\,d\tau \tag{16}$$

and is a measure of the rate at which stirring decreases the mean square gradient.

The other terms of (15) have been adequately discussed above, so that it remains only to discuss the integrand of (16). The velocity gradient is usually resolved into two components

$$\frac{\partial u_i}{\partial x_j} = u_{[ji]} + u_{(ji)}, \tag{17.1}$$

where

$$u_{[ji]} = \frac{1}{2}\left\{\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right\} \tag{17.2}$$

is recognized as being (except for a factor $-\frac{1}{2}$) the curl or vorticity of the motion, while the tensor

$$u_{[ji]} = \frac{1}{2}\left\{\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right\} \tag{17.3}$$

is called the deformation tensor and will be recognized from the familiar Stokes law of molecular viscosity.

It may be somewhat surprising to note that the integrand of (16) is independent of the curl, $u_{[ji]}$, but it can be shown that

$$\begin{aligned}\frac{\partial u_i}{\partial x_j}\frac{\partial \theta}{\partial x_i}\frac{\partial \theta}{\partial x_j} = u_{(11)}\left(\frac{\partial \theta}{\partial x_1}\right)^2 &+ 2u_{(12)}\frac{\partial \theta}{\partial x_1}\frac{\partial \theta}{\partial x_2} + 2u_{(13)}\frac{\partial \theta}{\partial x_1}\frac{\partial \theta}{\partial x_3}\\ &+ u_{(22)}\left(\frac{\partial \theta}{\partial x_2}\right)^2 + 2u_{(23)}\frac{\partial \theta}{\partial x_2}\frac{\partial \theta}{\partial x_3}\\ &+ u_{(33)}\left(\frac{\partial \theta}{\partial x_3}\right)^2.\end{aligned}$$

This is an important result, because it is known that curl or vorticity is very difficult to generate in a large mass of fluid whose coefficient of viscosity is small. If it should have appeared that the opposite were true, that S depended only on the curl, then the hypothesis under investigation would have to be rejected. Such a result would have made the process of mixing so strongly dependent on viscosity as to be untenable. The astonishment at the result arises from the popular confusion of curl or vorticity with eddy motion. The second example above shows that eddy motion is very effective in mixing, possibly more so than laminar motion. On the other hand, it is a fact that in laminar motion the curl cannot be zero, whereas it is easy to give an example of eddy motion with circular orbits for which the curl vanishes; e. g., $\omega(r) = 1/r$. Thus, there is no inconsistency between (16) and the two examples discussed earlier.

In the examples, and in the more general discussion of stirring, it was indicated that the effect of the velocity gradients is usually to increase the mean temperature or concentration gradients. Therefore, it might be expected that the quantity S is always negative. This cannot be proven; in the case of an incompressible fluid, the velocity must satisfy the equation

$$\nabla \cdot u = 0,$$

or

$$u_{(11)} + u_{(22)} + u_{(33)} = 0\,.$$

Hence, usually at least one of the three terms of this equation (say $u_{(11)}$) will be positive and at least one (say $u_{(22)}$) will be negative. Under these conditions, a gradient having the x_1 direction would make $S > 0$, while a gradient having the x_2 direction would make $S < 0$. Thus S may have either sign, and stirring may sometimes have the effect of decreasing the mean gradient. This situation will be clarified by a re-examination of the first example.

MORE DETAILED EXAMINATION OF MIXING BY LAMINAR FLOW

The example of stirring by laminar motion is simple enough so that fairly detailed calculations can be made. If the temperature (or concentration) at time $t = 0$ is

$$\theta = ax + by + c\,, \tag{18}$$

it is required to find a solution of

$$\frac{\partial\theta}{\partial t} + u(y)\frac{\partial\theta}{\partial x} = \kappa\,\nabla^2\theta\,, \tag{19}$$

which also reduces to (18) when t vanishes.

When $\kappa = 0$, this solution is unique, being

$$\theta \equiv \theta_o = a[x - u(y)t] + by + c\,. \tag{20}$$

This represents the effect of advection or stirring. The component of the gradient in the y-direction is

$$\frac{\partial\theta}{\partial y} = b - at\frac{du}{dy}\,.$$

For small values of t, its magnitude may either increase or decrease, depending on the relation between a, b and du/dy. For sufficiently large values of t, however, the second term must dominate, and hence the magnitude of the gradient will ultimately increase.

This illustrates two general propositions. It has already been noted above that advection may either increase or decrease the gradients. The new proposition is that, ultimately, advection usually increases the magnitude of the gradient. This is not without exception. In the textbook example of small surface waves, the motion is everywhere periodic and everywhere has the same period. Consequently, the changes caused by advection must also be periodic and they therefore constitute an exception to the rule. Actual surface waves, however, are not periodic, and their advective effects presumably conform to the rule. Because of these exceptional cases, a general proof of the proposition will be difficult and will be analogous to the proof of the ergodic theorem.[2]

Turning next to the case $\kappa \neq 0$, one may set

$$\theta = \theta_o + \vartheta\,, \tag{21}$$

---

[2] G. D. Birkhoff. Proof of a recurrence theorem for strongly transitive systems. Proc. nat. Acad. Sci., *17:* 650–655 (1931).

where ϑ is determined by the equation

$$\frac{\partial\vartheta}{\partial t} + u(y)\frac{\partial\vartheta}{\partial x} - \kappa\,\Delta^2\vartheta = -\,a\kappa t\frac{d^2u}{dy^2}. \tag{22}$$

At $t = 0$, one must require $\vartheta = 0$; it is characteristic that this requirement does not uniquely determine ϑ. Physically this means that conditions at the boundary of the medium are important: the sun may rise at some time after $t = 0$ and cause warming, etc.

However, special solutions of (22) can easily be obtained. One, which is independent of x, is the series

$$\vartheta = -\,a\sum_{m=2}^{\infty}\frac{\kappa^{m-1}t^m}{m!}\left(\frac{d}{dy}\right)^{2m-2}u(y)\,, \tag{23}$$

which can be considered as the solution that is applicable in the absence of such external effects as were just mentioned. It therefore determines the mixing process that is induced by laminar flow.

Since the series for ϑ begins with a term in t^2, it follows that, for sufficiently small values of t, the effect of mixing will be negligible compared to the advective effect which is measured by the term $-\,aut$ in (20). The time, t_c, at which the first term of the series for ϑ becomes equal to the advective term, is given by

$$t_c = 2u\Big/\left(\kappa\frac{d^2u}{dy^2}\right)$$

and is a function of y.

Since $u/(d^2u/dy^2)$ has the dimensions of an area, it may be set equal to l^2, where l is a length that measures the homogeneity of the velocity distribution. Hence

$$t_c = 2l^2/\kappa. \tag{24}$$

For water, the thermal diffusion constant is $\kappa = 1.4 \times 10^{-3}$ cm²/sec., while for the diffusion of dissolved salts, $\kappa = 2 \times 10^{-5}$ cm²/sec. This leads to the following values for t_c:

MIXING TIMES

l	thermal	salt
1 m.	168 d.	32 yr.
1 cm.	24 min.	28 hr.
0.1 mm.	0.15 sec.	10 sec.

These values suggest that the saline microstructure of the sea may be much more pronounced than the thermal microstructure, since diffusion can scarcely counteract the effect of mixing.

The solution for ϑ in the case of a circular eddy is more complicated but should be derivable without excessive difficulty.

THE EFFECT OF VISCOSITY

The two examples considered above are both steady states of flow, and it is unlikely that steady motion will persist over the relatively long time interval t_c. In particular, it can be shown that viscosity will usually modify the motion considerably in such a length of time, unless it is balanced by a suitable pressure distribution.

The viscous effects are governed by equations that are very analogous to (10), except that κ is replaced by the kinematic viscosity n. For water n/κ (Prandtl's number) has the value 7.5. Consequently, the viscous effects will become appreciable in a time that is less than that required for thermal mixing. They will, in general, tend to increase l and hence slow down the mixing process. However, this tendency may be counteracted by pressure gradients or wind stresses.

CONCLUSIONS

1. In all except certain very special cases, advection alone will ultimately increase the mean value of any initial gradient.
2. This effect of advection is appropriately called stirring.
3. Stirring is independent of the vorticity of the motion and can occur even if the motion is not turbulent. It may be the explanation of the high gradients shown by some bathythermograms.
4. The effect of conduction or diffusion is to decrease the mean value of the gradient.
5. This is appropriately called mixing.
6. Ordinarily, the early stages of a process in which both stirring and mixing occur will be dominated by the advective processes.
7. These may so increase the mean gradient that the mixing process will ultimately dominate over the stirring process.
8. Viscosity, if not counteracted by other factors, tends to stop the stirring process before an appreciable amount of mixing can occur.

NOTES ON A THEORY OF THE THERMOCLINE[1]

By

WALTER H. MUNK[2] AND ERNEST R. ANDERSON

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

The very large number of bathythermograms which have been taken during the last few years have established the essential features of the temperature structure in the upper layers of the ocean. When these layers are heated during a period of calm, the bathythermograms show a marked temperature gradient all the way to the surface, as shown in Fig. 1. If subsequently a strong wind starts to blow, the upper layers are stirred until an almost homogeneous layer is formed, bounded beneath by a region of marked temperature gradient, the thermocline. The transition from the homogeneous layer to the thermocline is usually sudden. If the wind increases in intensity the thermocline moves downward, but the characteristic *shape* of the temperature-depth curve remains essentially unchanged.

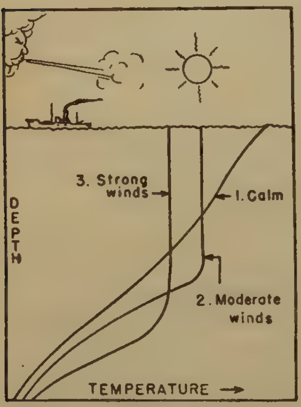


Figure 1. Schematic presentation of the development of a wind stirred layer and thermocline.

It has been customary to interpret the wind-stirring effect in terms of Ekman's famous theory for wind driven currents on a rotating globe (Ekman, 1905; Sverdrup, *et al.*, 1942: 492–494). Ekman assumes *homogeneous* water and a constant value[3] for the eddy viscosity, and obtains for a solution current vectors whose azimuths vary linearly and whose magnitudes decrease exponentially with depth from the very surface downward. Therefore, the Ekman spiral does not contain the concept of a homogeneous upper layer with a well defined lower boundary. Nevertheless, this concept has often been brought in by means of the artifice of identifying the thickness of the mixed layer with a depth where the current equals $e^{-\pi}$ times the surface current and is directed opposite to it.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 382. This work represents the results of research carried out for the Office of Naval Research, the Hydrographic Office, and the Bureau of Ships of the Navy Department, under contract with the University of California.

[2] Institute of Geophysics and Scripps Institution of Oceanography, University of California.

[3] Ekman (1905) also gives a solution for the case where the eddy viscosity is proportional to the shear. This does not modify the essential meaning of our remarks.

The difficulty can be resolved, at least qualitatively, by taking into account the effect of the density stratification on the scale of turbulence. Ekman (1905) stated:

> It is obvious that [the eddy coefficient] cannot generally be regarded as a constant when the density of the water is not uniform within the region considered. For [the eddy coefficient] will be greater within the layers of uniform density and comparatively small within the transition-layers where the formation of vortices must be much reduced owing to the differences of density.

The reduction in the vertical eddy conductivity in the thermocline leads to high resistance to heat flux from above, with ". . . the paradoxical result that the water-masses below . . . are the more 'protected' against heating the stronger the heating . . ." (Helland-Hansen, 1930: 44). Surface heating therefore tends toward sharpening the existing temperature gradient in the thermocline and further reducing the turbulence. At the same time, the reduction in turbulence permits relatively easy slippage of water layers with respect to one another, and the resulting shear in the thermocline represents a source for producing turbulent energy. The ultimate outcome must be a steady-state distribution which represents a balance between the stabilizing influence of temperature gradients and the unstabilizing influence of current gradients. It is this steady-state solution which is being sought.

The foregoing discussion indicates how intimately the problems of current and of temperature distributions are related. Mathematically they are linked by the coefficients of eddy viscosity and conductivity, both of which are functions of the gradients of temperature *and* current. For the eddy viscosity such a function has been introduced by Rossby and Montgomery (1935); for the eddy conductivity a new relationship will be derived. These functions will be substituted into the differential equations governing the distributions of temperature and current, and the differential equations solved *simultaneously*.

Ekman was well aware of the desirability of incorporating these thermodynamic considerations into his dynamic treatment of ocean currents. He wrote that ". . . this is however a complication which cannot be taken into account in the general theory . . ." We have been fortunate to have at our disposal the new Differential Analyzer of the University of California. With this computer, numerical solutions have been found for 21 examples; 14 examples correspond to actual conditions for which complete observational material was available, and which were in agreement with the assumptions underlying the theory. The most restrictive of these assumptions is that the net heat flux be directed downward and that there be no convective stirring such as might be induced by surface cooling due to evaporation.

EQUATIONS OF TURBULENT FLUX UNDER STABLE CONDITIONS

The downward flux, per unit area, of momentum, of heat and of salt, due to turbulent processes, equal respectively

$$F_M = -A_V V', \quad F_T = -c_p A_T T', \quad F_S = -A_S S', \qquad (1\text{ a, b, c})$$

where c_p is the specific heat at constant pressure, A_V, A_T, and A_S the dynamic eddy coefficients (dimensions mass length^{-1} time^{-1}) of viscosity, conductivity and diffusivity, respectively. The accents ($'$) denote differentiations with respect to depth:

$$V' = \partial V/\partial z, \qquad T' = \partial T/\partial z, \qquad S' = \partial S/\partial z, \qquad (2)$$

positive z being directed downward. It is customary to refer to V' as shear and to

$$\tau = -F_M \qquad (3)$$

as stress.

The coefficients A_V, A_T and A_S are not physical constants of the fluid but depend upon the state of its motion and may vary with time and position. Two factors tend to reduce the values of the turbulence coefficients: boundaries and stable stratification. Since the effect of the boundary is limited to the upper few meters of the ocean, we shall be concerned only with stability.

Let the subscript "o" denote the case of *neutral stability.* The ratio A/A_o for any one of the three coefficients in equation (1) is assumed to depend explicitly on stability, shear and gravity:

$$A/A_o = f(E, V', g), \qquad (4)$$

where the stability E depends upon the temperature and salinity distributions

$$E = E(T', S', T, S), \qquad (5)$$

and these distributions vary with depth and time. The theory of the wind-stirred layer is contained in the solution $T(z)$ of equations (4), (5) and the three equations (1).

Stability, Temperature and Salinity Gradients. It is convenient, whenever possible, to express the stability as a function of temperature gradient or salinity gradient only:

$$E = a\,T' \qquad \text{or} \qquad E = b\,S'. \tag{6}$$

To examine the implications of equation (6) and to determine the coefficients a and b, we make use of equation XII, 12–14, in Sverdrup, *et al.* (1942), according to which

$$E = \frac{1}{\rho}\frac{\partial \rho}{\partial z} \doteq 10^{-3}\frac{\partial \sigma_t}{\partial z} \tag{7}$$

except at great depth. Expanding by partial differentiation one finds

$$E = 10^{-3}\left(\frac{\partial \sigma_t}{\partial T}T' + \frac{\partial \sigma_t}{\partial S}S'\right) = 10^{-3}\left(\frac{\partial \sigma_t}{\partial T} + \phi\frac{\partial \sigma_t}{\partial S}\right)T', \tag{8}$$

or

$$E = 10^{-3}\left(\frac{1}{\phi}\frac{\partial \sigma_t}{\partial T} + \frac{\partial \sigma_t}{\partial S}\right)S', \tag{9}$$

where, provided S and T are functions of z only,

$$\phi = \frac{S'}{T'} = \frac{dS}{dT} \tag{10}$$

is the slope of the line denoting the T-S relationship of the water mass. Comparison between equations (6) and (8) or (9) gives

$$a = 10^{-3}\left(\frac{\partial \sigma_t}{\partial T} + \phi\frac{\partial \sigma_t}{\partial S}\right), \quad b = 10^{-3}\left(\frac{1}{\phi}\frac{\partial \sigma_t}{\partial T} + \frac{\partial \sigma_t}{\partial S}\right). \tag{11}$$

To simplify computations, the parameters a or b will be considered independent of depth in subsequent integrations. This involves the assumptions that over the range of integrations the T-S relationship is linear, with variations in temperature and salinity small. The errors introduced by these assumptions are generally small compared to those inherent in Ekman's assumption of a constant eddy viscosity.

TABLE I. VALUES OF $\partial\sigma_t/\partial T$

T \ S	33°/oo	35°/oo	37°/oo
0 °C	−.045	−.055	−.060
10 °C	−.170	−.172	−.175
20 °C	−.260	−.262	−.265

TABLE II. VALUES OF $\partial\sigma_t/\partial S$

T \ S	33°/oo	35°/oo	37°/oo
0 °C	.806	.806	.807
10 °C	.779	.780	.780
20 °C	.760	.761	.763

Effect of Stability and Shear on Eddy Coefficients. According to the Π-theorem in dimensional analysis, the quantities E, V', and g, on which the ratio of eddy coefficients was assumed to depend (equation 4), can be combined into only one type of nondimensional number, the Richardson number r (Brunt, 1939: 237):

$$r = g\,E/(V')^2. \tag{12}$$

In selecting a particular functional relationship we are guided by the limiting conditions

$$A \rightarrow A_o \text{ for } r \rightarrow 0, \tag{13}$$

$$A \rightarrow 0 \text{ for } r \rightarrow \infty. \tag{14}$$

Equation (14) states that for very high stability the flux of momentum heat and salt takes place by molecular processes only.

A simple set of equations which satisfy (13) and (14) is

$$A_V = A_o\,(1 + \beta_V\, r)^{-n_V}, \tag{15}$$

$$A_T = A_o\,(1 + \beta_T\, r)^{-n_T} = A_S, \tag{16}$$

where the eddy coefficients for *neutral* stability

$$A_{V,\,o} = A_{T,\,o} = A_{S,\,o} = A_o \tag{17}$$

have been assumed equal,[4] and where β_V, n_V, β_T, n_T are positive constants, not yet determined. The coefficients pertaining to heat flux and salinity diffusion are assumed to equal one another but to differ from the coefficient pertaining to momentum flux (see below).

Equation (15) is a relationship derived by Rossby and Montgomery (1935). It will be shown that the constants in (16) can be evaluated from known relationships, and that the resulting equations are in agreement with observations.

DETERMINATION OF CONSTANTS

Application of Investigations by Rossby and Montgomery and by Sverdrup. From energy considerations, Rossby and Montgomery (1935) derived an equation identical in form to equation (15), with $n_V = \frac{1}{2}$. Sverdrup (1936), on the basis of his measurements over a snow field at Spitzbergen, estimates β_V at 10 to 13, and probably nearer the lower limit. We shall adopt the value $\beta_V = 10$.

---

[4] "In recent years a great deal of work on the relationship between A_S and A_V has been done in the field of Aeronautics. The assumption of Reynold and Prandtl that $A_S = A_V$ has been confirmed in all cases when the fluid is homogeneous" (Taylor, 1931, notation ours).

Jacobsen's Investigation. From his extensive analysis of salinity and current observations in Randers Fjord and Schultz Ground, Jacobsen (1913) postulates that under stable conditions the turbulent elements give off their momentum rapidly to their surroundings, whereas other properties, such as temperature and salinity, are exchanged slowly. The elements are therefore moved to new surroundings by gravitational forces before equalization can take place. As a consequence, ". . . the effect of stability on turbulence is two-fold. In the first place, the turbulence is reduced, leading to smaller values of the eddy viscosity, and, in the second place, the type of turbulence is altered in such a manner that the accompanying eddy diffusivity becomes smaller than the eddy viscosity" (Sverdrup, *et al.*, 1942: 477).

The first conditions, $A_V \leqq A_0$, is already satisfied by equation (15); the implication of the other condition,

$$X_j \equiv A_S/A_V \leqq 1, \tag{18}$$

remains to be investigated. Substituting from equations (15) and (16) gives

$$X_j = (1 + \beta_V r)^{n_V} (1 + m \beta_V r)^{-n_T}, \tag{19}$$

where

$$m = \beta_T/\beta_V. \tag{20}$$

For $r = 0$ one finds

$$X_j = 1, \; d X_j / d r = - \beta_V n_T (m - n_V/n_T), \tag{21}$$

so that

$$m > n_V/n_T, \tag{22a}$$

in order for Jacobsen's condition (18) to be satisfied (Fig. 2). The special case

$$m = n_V/n_T, \tag{22b}$$

for which X_j approaches 1 asymptotically as r approaches zero, seems more likely from a physical point of view than the general case for which $d X_j / d r$ has a finite value at $r = 0$.

For very large values of r,

$$X_j \doteq m^{-n_T} (\beta_V r)^{n_V - n_T}, \tag{23}$$

and

$$n_T > n_V, \tag{24}$$

in order for (18) to be satisfied. Equations (22) and (24) impose certain restrictions upon the numerical values of n_T and β_T.

Taylor's Investigation. Further restrictions are imposed by an entirely different condition proposed by Taylor (1931):

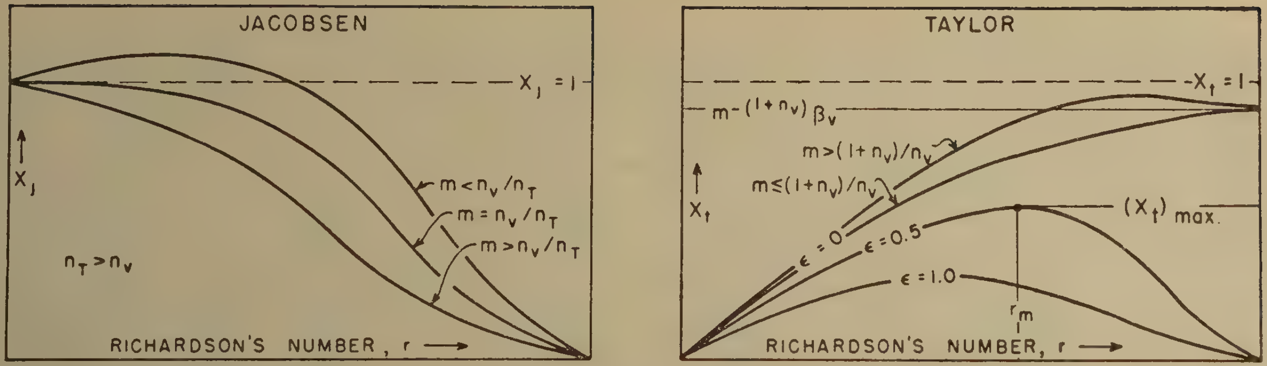

JACOBSEN
$X_j = 1$
X_j
$m < n_v/n_T$
$m = n_v/n_T$
$m > n_v/n_T$
$n_T > n_v$
RICHARDSON'S NUMBER, r
TAYLOR
$X_t = 1$
$m-(1+n_v)\beta_v$
$m > (1+n_v)/n_v$
$m \leq (1+n_v)/n_v$
X_t
(X_t) max.
$\epsilon = 0$
$\epsilon = 0.5$
$\epsilon = 1.0$
r_m
RICHARDSON'S NUMBER, r

Figure 2. Limitations imposed by Jacobsen's and by Taylor's conditions upon the constants in equations (15) and (16). See text.

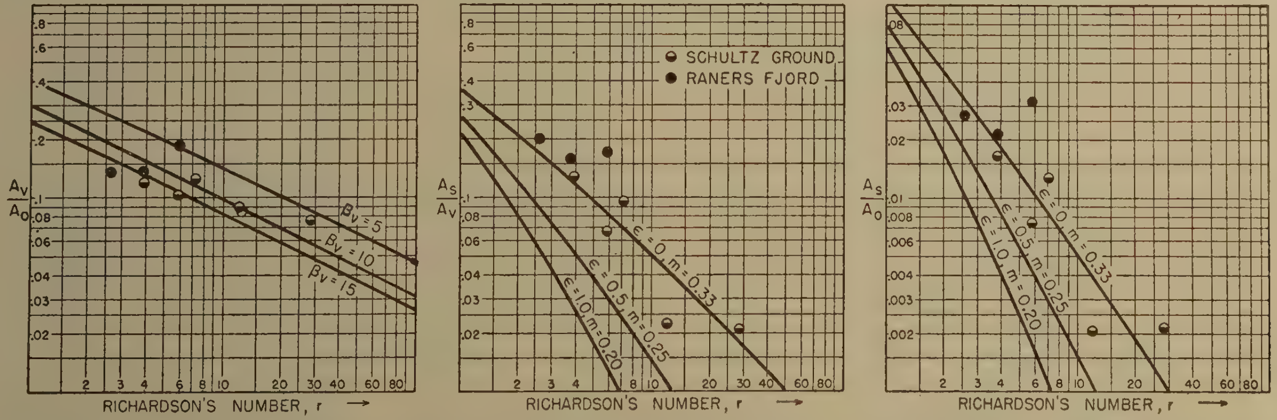


Figure 3. Comparison of theoretical relationships between eddy coefficients and Richardson's number with that derived from Jacobsen's observations of current and salinity. The left part of the figure confirms the value of β_V obtained by Sverdrup from meteorological measurements over a snow field. The central and right parts are consistent with values of ϵ and m derived from physical considerations.

In the case of a stratified fluid with density increasing downward any disturbance to steady laminar flow necessarily involves an increase in gravitational potential energy. It is this increase which causes infinitely small disturbances of a fluid at rest to be stable. In cases of turbulent flow the rate at which energy flows into a layer of the fluid owing to the action of the fluid above and below it must be greater than the rate at which gravitational energy increases in the same portion of the fluid.

Taylor shows that the fraction of available energy which is used to increase gravitational potential energy equals $r\,A_T/A_V$, and his argument requires therefore that

$$X_t \equiv r\,A_T/A_V \leqslant 1. \tag{25}$$

Again returning to (15) and (16) one finds

$$X_t = r\,(1 + \beta_V\,r)^{n_V}\,(1 + m\,\beta_V\,r)^{-n_T}. \tag{26}$$

For very small values of r, equation (26) reduces to $X_t \doteq r$, independent of β_T and n_T (Fig. 2). For large values of r,

$$X_t \doteq m^{-n_T}\,\beta_V^{\,n_V - n_T}\,r^{1+n_V-n_T}, \tag{27}$$

so that $n_T \geqslant 1 + n_V$, or

$$n_T = 1 + n_V + \varepsilon, \quad \varepsilon \geqslant 0, \tag{28}$$

in order for Taylor's condition (25) to be satisfied.

In general X_t reaches a maximum, to be designated by $(X_t)_{\max}$, for $r = r_m$ (Fig. 2), where r_m is a root of the equation

$$m\,\varepsilon\,\beta_V^2\,r_m^2 + \beta_V\,[m\,(n_V + \varepsilon) - (1 + n_V)]\,r_m - 1 = 0. \tag{29}$$

As long as Taylor's condition holds for $(X_t)_{\max}$ it certainly must hold for any other values of X_t, and we may express equation (25) in the more definite form

$$(X_t)_{\max} \leqslant 1. \tag{30}$$

For the special case

$$\varepsilon = 0, \text{ and } m \leqslant (1 + n_V)/n_V, \tag{31}$$

a true maximum does not exist, but X_t remains bounded (Fig. 2). The term $(X_t)_{\max}$ in equation (30) is then interpreted as $m^{-(1+n_V)}\,\beta_V^{-1}$, the value of X_t at $r = \infty$. The special conditions stated in equation (31) appear more reasonable than any of the other possible combinations, as these other combinations require that, with increasing stability, the fraction X_t of the available energy going into potential energy should first increase and then decrease.

Equations (28), (29), and (30) define a minimum value for m for any fixed value of ε. Limitations imposed by Jacobsen's and by Taylor's conditions on the constants n_T and β_T are summarized in Table III.

Taylor's condition imposes the more severe restriction on n_T, whereas Jacobsen's condition imposes the more severe restriction on m, or β_T.

TABLE III. Limits to Constants Imposed by Jacobsen's and by Taylor's Conditions

Constant	n_V	n_T		ε	n_T	m	
Source	Rossby	Jacobsen	Taylor	Assumed	——	Jacobsen	Taylor
Equation	—	(24)	(28)	—	(28)	(22a)	(28) (29) (30)
Values	0.5	$\geqslant 0.5$	$\geqslant 1.5$	0.0*	1.5	$\geqslant 0.33$†	$\geqslant 0.22$
—	—	—	—	0.5	2.0	$\geqslant 0.25$	$\geqslant 0.10$
—	—	—	—	1.0	2.5	$\geqslant 0.20$	$\geqslant 0.07$

* Preferable for reasons stated following equation (31).

† $m = 0.33$ preferable to $m > 0.33$ for reasons stated following equation (22b).

For reasons referred to in the footnotes to Table III, the combination

$$\varepsilon = 0, \; n_T = 1.5, \; m = 0.33 \tag{32}$$

appears to be the most reasonable. Substituting from (15) and (16) gives

$$\begin{aligned} A_V &= A_o \, (1 + \beta_V \, r)^{-1/2}, \; \beta_V = 10 \\ A_S = A_T &= A_o \, (1 + \beta_T \, r)^{-3/2}, \; \beta_T = \beta_V/3 = 3.33 \, . \end{aligned} \tag{34}$$

Some Numerical Values. Taylor (1931) has used Jacobsen's observations to test the validity of his criterion, and he finds condition (25) satisfied. The same observations have been used in Fig. 3, eliminating, however, the observation at 15 meters, where the shear shows a sudden change from neighboring depths which is inconsistent with the general trend.

The left part of Fig. 3 indicates the validity of equation (15), using $\beta_V = 10$. It has already been stated that this value was derived from meteorological observations; its applicability to oceanographic data may indicate that β_V represents a universal constant. The curves in the middle and right parts of Fig. 3 were computed for three values of ε and the corresponding *minimum* values of m (second column from right in Table III.) Increasing m shifts the curves toward the left away from the observations. The empirical data are consistent with the choice of values proposed in equation (32) on the basis of physical reasoning.

THE DISTRIBUTIONS OF VELOCITY AND TEMPERATURE

General Solution. The problem consists of solving simultaneously the five equations

$$\tau = A_V V', \qquad F_T = -c_p A_T T', \tag{1a, b}$$

$$A_V = A_o (1 + \beta_V r)^{-1/2}, \quad A_T = A_o (1 + \beta_T r)^{-3/2} \tag{33, 34}$$

$$r = g a T' (V')^{-2} \tag{35}$$

in the five unknowns V, T, A_V, A_T and r. Solving for r gives

$$\frac{g(-a) F_T A_o}{c_p \tau^2} = r (1 + \beta_V r)(1 + \beta_T r)^{-3/2}, \tag{36}$$

where, in general, a, F_T, A_o, and τ are all functions of z, and these functions must be specified in order that solutions $r = r(z)$ can be found. It will be convenient to introduce functions K and R so that equation (36) becomes

$$K(z) = R(r). \tag{37}$$

The distributions of velocity and temperature are best expressed in terms of r;

$$\frac{V'}{V_o'} = \frac{(1 + \beta_T r)^{3/4}}{r^{1/2}}, \quad V_o' = \frac{g(-a) F_T}{c_p A_o}, \tag{38}$$

$$\frac{T'}{T_o'} = (1 + \beta_T r)^{3/2}, \quad T_o' = -\frac{F_T}{c_p A_o}. \tag{39}$$

The above relationships are plotted in Fig. 4.

Depth and Sharpness of Thermocline. Before taking up the special case of the Ekman spiral, a few quite general remarks concerning the character of the thermocline can be made. These remarks are applicable to transient as well as steady-state conditions.

Two features which are usually prominent on bathythermograms are: (a) the maximum temperature gradient, $T' = T_{\max}'$ at depth $z = z_g$; (b) the "knee," $T'' = T_{\max}''$ at depth $z = z_c$. As a simple case, assume T_o' independent of z. It can be seen from Fig. 4 that in order for the temperature gradient to have a maximum value, r and hence $K(z)$ must reach a maximum. Other factors remaining equal, *the stress must be at a minimum where the temperature gradient is at a maximum. The maximum temperature gradient will be the larger, the smaller the value of the minimum stress.*

At the depth z_c of maximum curvature $T''' = 0$, or, according to (39),

$$r'' + (\tfrac{1}{2}) \beta_T (1 + \beta_T r)^{-1} r'^2 = 0. \tag{40}$$

Combining this with the equations

$$K' = \frac{dR}{dr} r' , \tag{41}$$

$$K'' = \frac{d^2R}{dr^2} r'^2 + \frac{dR}{dr} r'' , \tag{42}$$

and substituting $\beta_T = \beta_V/3$, gives, after much manipulation,

$$K'' + QK'^2 = 0, \tag{43}$$

where

$$Q = \frac{2\,\beta_V\,(1 + \beta_T\,r)^{3/2}\,(\beta_V{}^2\,r^2 + 16\,\beta_V\,r - 15)}{(\beta_V{}^2\,r^2 + 11\,\beta_V\,r + 6)^2} \tag{44}$$

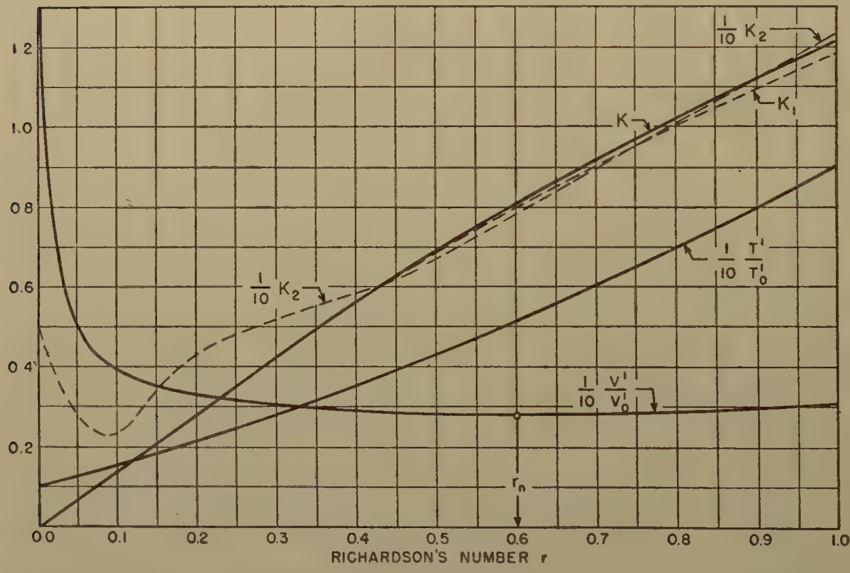


Figure 4. Plot of shear, temperature gradient, and the parameters K, K_1 and K_2 (see text) as functions of Richardson's number r. The shear reaches a minimum for $r = 0.6$.

is a function of r, or in view of (37), a function of K. The solution to (43) is

$$K_1 = C_1\,z + C_2, \tag{45}$$

where

$$K_1 = \int e^{-\int Q dK} dK$$

has been plotted in Fig. 4. Except for large values of r, $K_1 \doteq K$. *Therefore the depth* z_c *of maximum curvature in the* T(z) *curve equals approximately the depth of zero curvature in the* K(z) *curve.*

The maximum curvature can be computed from

$$T_{\max}'' = 3/2\ \beta_T\ (1 + \beta_T\ r)^{1/2}\ r'.$$

According to (40) and (45)

$$r' = \frac{K'}{R'} = \frac{C_1\, e^{\int Q dK}}{R'}$$

and

$$T_{\max}'' = K_1'\ (z_c)\ K_2\ (z_c), \tag{47}$$

where

$$K_2 = \frac{(3/2)\, \beta_T\, (1 + \beta_T\, r)^{1/2}\, e^{\int Q dK}}{R'} \tag{48}$$

is plotted in Fig. 4. Thus if $K(z)$ is known, the depth of the thermocline and the sharpness of the "knee" on the temperature-depth curve can readily be computed.

Minimum Shear Layer. The shear reaches a minimum value

$$V'_{\min} = \left[\frac{3\sqrt{3}}{2}\ \frac{g\ (-a)\ F_T\ \beta_T}{c_p\ A_o}\right]^{1/2} \tag{49}$$

when

$$r = r_n = 2/\beta_T = 0.6. \tag{50}$$

The corresponding values of A_V/A_o, A_T/A_o and K are 0.38, 0.19, and 0.81 respectively. The region of minimum shear occurs somewhat above the thermocline (Fig. 6). At greater depths the shear is larger because of the high stability which permits relatively easy slippage of water layers with respect to one another; at lesser depths the shear is larger because of the driving effect of the wind. Measurements to determine the possible existence of a minimum shear layer could serve as a critical test of this theory, especially as no corresponding feature is contained in Ekman's solutions.

The Modified Ekman Spiral. For steady-state conditions, with the frictional force balanced by the Coriolis force, the equations of motion take the form

$$\begin{aligned} \tau_x' &= \frac{d}{dz}\,(A_V\ V_x') = -\ C\ V_y \\ \tau_y' &= \frac{d}{dz}\,(A_V\ V_y') = C\ V_x\,, \end{aligned} \tag{51}$$

where

$$C = 2\omega \sin \varphi,$$

ω being the earth's rotational velocity, and φ the latitude, and where

$$\tau = \tau_x^2 + \tau_y^2, \qquad V' = V_x'^2 + V_y'^2, \qquad V = V_x^2 + V_y^2 \quad (52)$$

are the magnitudes of the stress, shear and velocity vectors, respectively. Equations (51) are identical in form with those solved by Ekman, the difference being that here A_V is a function of temperature gradient and shear, whereas Ekman treated A_V as constant, or a function of shear only.

Solutions were obtained by solving equations (1), (33), (34), (35) and (51) simultaneously on the differential analyzer. In order to have steady-state temperature distribution, the heat flux F_T was assumed independent of depth.

At the surface we set, with Ekman,

$$\tau_x = 0, \ \tau_y = -\tau_a \qquad \text{for } z = 0, \qquad (53)$$

where τ_a is the stress of the wind on the sea surface, assuming the wind to blow along the positive y-axis. At great depths the only reasonable boundary condition is the one prescribed by Ekman, namely that the velocity vanishes. This condition cannot be reconciled with the existence of a minimum value for the velocity gradient unless the heat flux vanishes at great depth (equation 49). It appears that the distribution of temperature and current cannot both be stationary at the same time.

Without the bottom boundary conditions the solutions to the equations are no longer uniquely determined. To make the problem definite, we shall assume the components of surface current to be given by Ekman's equations

$$V_x = V_y = \frac{\tau_a}{\sqrt{2\,\rho\,A_o\,C}} \text{ for } z = 0. \qquad (54)$$

Values of A_o and τ_a as functions of the wind speed were read off Fig. 5. The former are based on a comparison of Ekman's theory with observations (Sverdrup, *et al.*, 1942: 494), the latter on recent calculations[5] taking into account the roughness of the sea surface and the gustiness of the wind.

Fig. 6 shows the results of a sample calculation (run 3b, Table IV). Minimum shear is found at a depth of 15 meters, the maximum

---

[5] A discussion by Mr. T. Saur is in preparation.

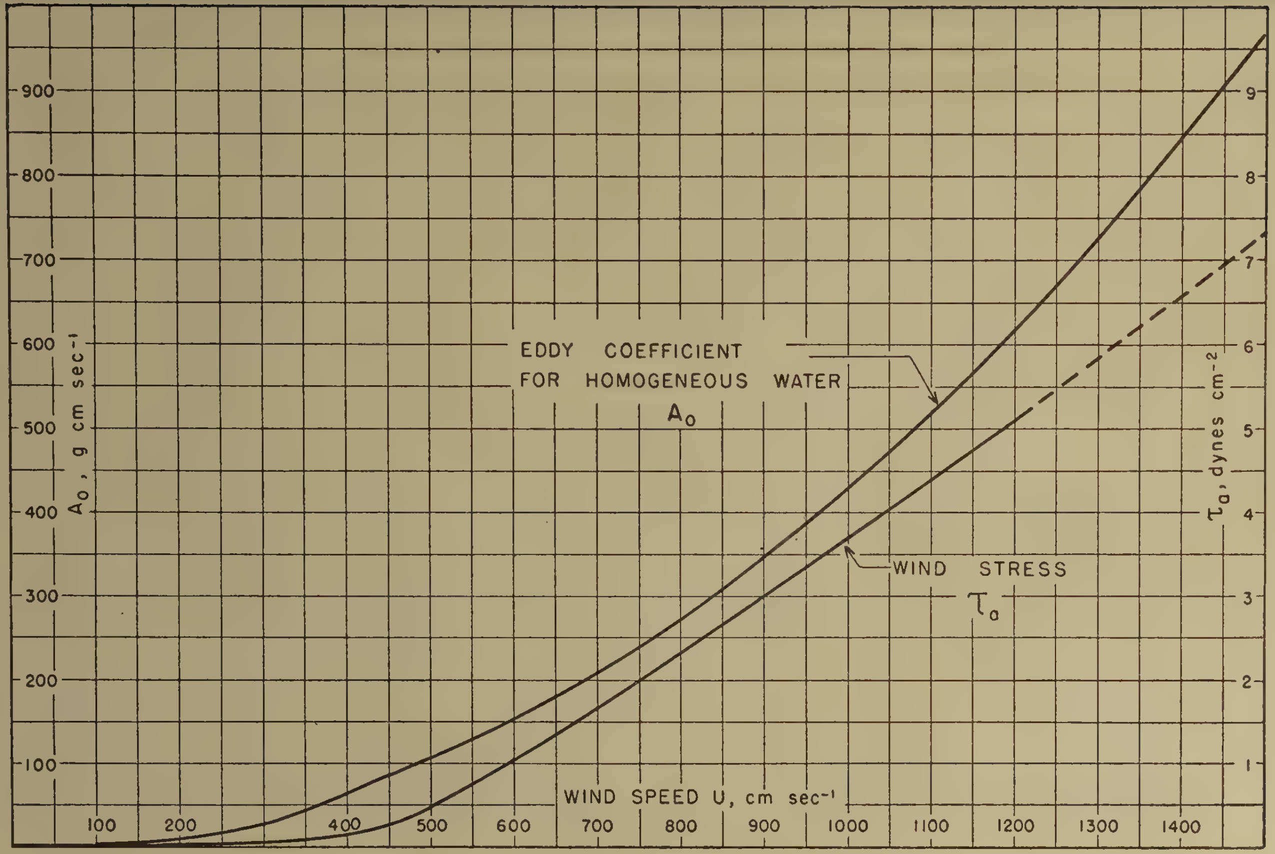


Figure 5. Relationships between wind speed at standard anemometer elevations, the wind stress per unit area of sea surface, and the eddy coefficient for homogeneous water.

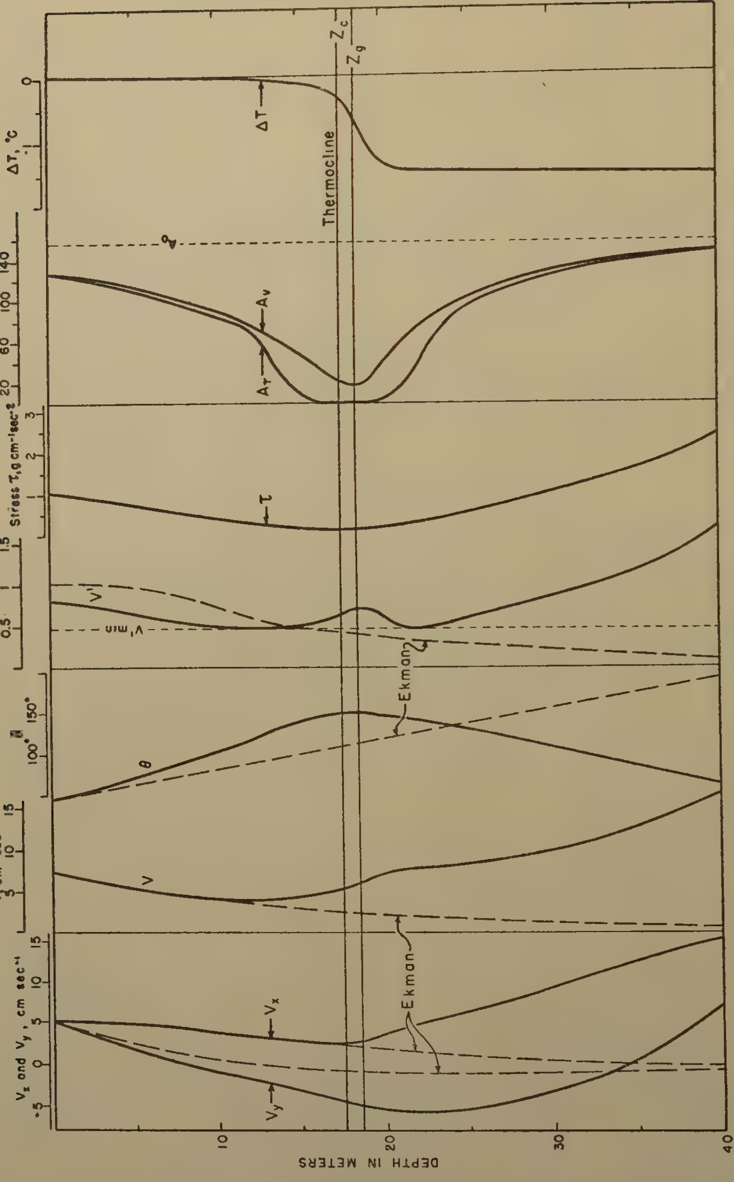


Figure 6. Results of computation on Differential Analyzer for Run 3b, Table IV.

curvature in the temperature depth curve at 17.5 meters, and the maximum temperature gradient at 18.5 meters. At the latter depth the stress reaches a minimum as predicted by the general theory.

Because of the simplifying assumption that F_T is a constant, the solution becomes increasingly inadequate the greater the depth.

Variation in Parameters. All parameters entering the differential equations and, according to our theory, the distributions of temperature and current in the upper layers of the oceans, depend upon four factors: the wind speed U, heat flux F_T, latitude φ, and the T-S correlation factor a. Some of these factors enter the equations more than once; U enters in three places. To determine the role played by each factor, computations have been carried out for typical conditions, varying one factor at a time.

The first five lines in Table IV show the results of the computations. Doubling the wind speed lowers the thermocline from 22 to 61 meters! Other factors have a much smaller effect. Doubling the latitude raises the thermocline by 40%, doubling the heat flux by 20%, doubling a by only 15%.

The next two lines in Table IV indicate that the computed depth of the thermocline does not depend critically upon the assumptions made in equation (54) as to the relative direction and velocity of the surface current.

COMPARISON WITH OBSERVATIONS

In selecting specific cases to be used for comparison with theory the following considerations were involved: (1) bathythermograph observations must be adequate; (2) evaporation relatively small; (3) heat flux positive; and (4) advection small, *i. e.*, no strong currents. The depth of thermocline was computed for: (1) a relatively clear-cut situation observed from a weather ship; (2) a broad strip in mid-Pacific, for which the computed and observed changes of thermocline depth with *latitude* could be compared; and (3) a lake at various times of the year, for which the computed and observed changes with *season* could be compared. Results are summarized in Table IV.

The weather ship was located at Lat. 49° 50′ N, Long. 145° 30′ W; the time of observation was September 28–29, 1943. Five bathythermographs taken during the 20-hour period were averaged to minimize the effect of internal waves. During this time interval there was condensation on the sea surface. The heat flux was computed from the general heat budget, following the method described on pp. 100–124 of Sverdrup, *et al.* (1942). The constant a was obtained from a hydrographic station occupied by the USS Bushnell on August 19, 1934, at Lat. 49° 29′ N, Long. 174° 31′ W. The observed depth of maximum curvature of the bathythermogram

TABLE IV. SUMMARY OF COMPUTATIONS ON DIFFERENTIAL ANALYZER

Run Number	*Wind Speed*	*Latitude*	*Heat Flux*	*T–S Correlation*	*Surface Speed*	*Wind Angle*	*Wind Stress*	*Eddy Viscosity*	*Depth z in Meters*		
	U (cm sec^{-1})	φ	$10^3 F_T$ (cal cm^{-1} sec^{-1})	$10^4 a$ (°C)$^{-1}$	$(v)_{z=0}$ (cm sec^{-1})	θ	τ (dynes cm^{-2})	A_0 (g cm^{-1} sec^{-1})	*Minimum Shear*	*Thermocline (T''' = 0) Computed*	*Observed*
	Variation of Parameters										
1h	600	30° N.	2.00	2.00	9.60	45.0°	1.02	155.0	19.6	22.0	—
2a	1200	30	2.00	2.00	24.30	45.0	5.15	619.0	58.9	61.0	—
3a	600	60	2.00	2.00	9.60	45.0	1.02	155.0	—	14.0	—
4	600	30	1.00	2.00	9.60	45.0	1.02	155.0	22.8	26.5	—
5	600	30	2.00	4.00	9.60	45.0	1.02	155.0	14.7	19.0	—
1j	600	30	2.00	2.00	9.60	48.4	1.02	155.0	18.1	21.5	—
3b	600	60	2.00	2.00	7.30	45.0	1.02	155.0	13.5	17.5	—
	Weather Ship (September, 1943)										
6	630	49	3.89	2.18	8.75	45.0	1.20	171.0	13.8	16.2	38.0
	Mid-Pacific NW–SE Strip (Average Summer)										
7	617	15	0.49	2.85	14.25	45.0	1.12	164.0	38.6	44.0	76.0
17	505	21	0.76	2.85	6.58	45.0	0.50	110.0	15.5	20.0	43.0
13	530	24	0.97	2.80	7.45	45.0	0.63	121.0	15.0	19.0	48.0
18	462	28	1.32	2.65	3.91	45.0	0.31	91.8	6.5	11.0	33.0
8	365	32	1.73	2.72	2.59	45.0	0.16	49.6	—	5.6	19.0
15	440	38	2.04	2.80	2.78	45.0	0.24	83.2	—	6.3	14.0
19	440	43	2.10	2.60	2.64	45.0	0.24	83.2	0.5	6.0	17.0
9	542	48	2.14	2.43	5.90	45.0	0.69	126.0	10.0	13.0	27.0
	Sweetwater Lake (April, May, June, July, August)										
10	350	32	1.18	1.89	2.58	45.0	0.15	43.7	3.5	7.2	6.5
14	330	32	2.28	2.07	2.63	45.0	0.14	36.6	—	4.5	6.5
11	365	32	2.00	1.33	2.59	45.0	0.16	49.6	2.3	7.5	5.5
16	325	32	2.20	2.23	2.50	45.0	0.13	35.0	—	3.0	4.0
12	300	32	1.18	1.83	2.39	45.0	0.11	27.5	2.3	4.5	4.0

trace was at 38 meters, the computed depth at 16 meters, giving a ratio of 0.43.

Computation for a mid-Pacific NW-SE strip was made for the summer season, during which evaporation is relatively small. The region is shown in the inset of Fig. 7. Values of *a* were computed from selected USS Bushnell hydrographic stations. Wind speeds were obtained from pilot charts of the North Pacific. Values of heat flux for each five-degree square were kindly put at our disposal by Dr. W. C. Jacobs. These were computed on the basis of energy budget considerations, and they may be considerably in error because of uncertainties involved in some of the data. The observed depths of the thermocline were taken from a compilation by Mr. E. C. LaFond, using all data available prior to January 1945. For the extreme northern and southern portions of the area, the averages were adjusted for recent observations. Fig. 7 shows the variation with latitude of the basic parameters and of the observed and computed depths of thermocline. The computed depth is much too shallow, bearing roughly the same ratio to the observed depth as was found for the

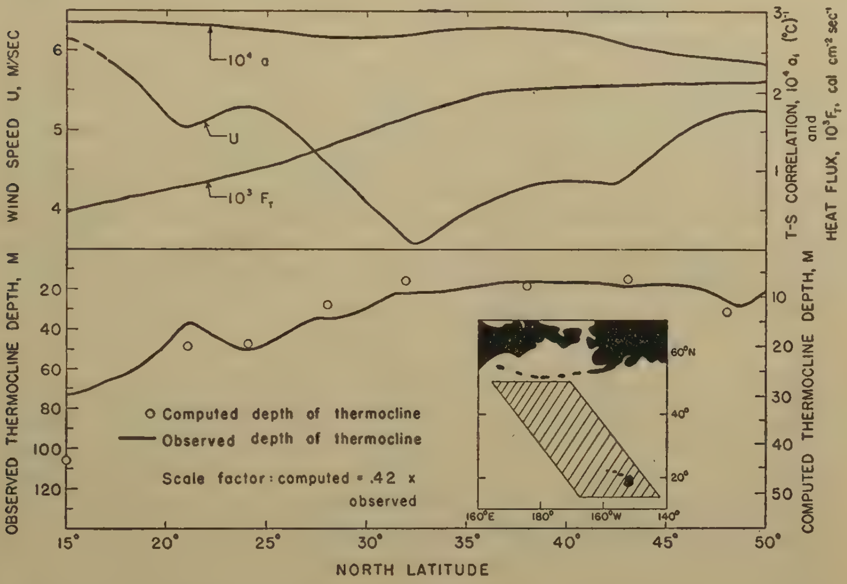


Figure 7. Comparison between computed and observed depths of thermocline as functions of latitude for the broad strip in the mid-Pacific shown in inset. The upper part gives the variations of the basic parameters entering the calculations.

weather ship. Except for the southernmost station, the percentage *variations* in the depth of thermocline are similar for the computed and observed values.

Observations for Sweetwater Lake, California, taken during 1944 and 1945, have kindly been placed at our disposal by Mr. B. E. Holtsmark. Temperature observations were taken from a barge in water of about 50-foot depth. No wind observations were taken at the time, and it has been necessary to use wind data taken at Lindberg Field, San Diego, located approximately 12 miles from Sweetwater Lake. Values of a were obtained from standard tables of density of fresh water at various temperatures. The heat flux, which was computed by Mr. Holtsmark, using the energy budget method, was positive during the months under consideration. The agreement between the computed and observed depths of thermocline is fairly good. In addition, computations revealed an appreciable temperature gradient from the very surface downward, a feature in agreement with observations.

CONCLUSIONS

In view of the difficulty experienced in selecting conditions under which heat loss by evaporation is relatively small compared to heat gain by radiation, it would appear that convective stirring is as important as wind stirring in many, and perhaps most, cases. Therefore, it would be preferable to refer to the upper mixed layer by a less implicating term than "wind stirred layer."

The present model is greatly over-simplified. The effects of convective stirring, and of the convergence and divergence of the horizontal current field, must exert predominant influences in certain regions. If the heat flux had been assumed to decrease quasi-exponentially with depth, to take into account the absorption of radiation, more consistent results would have been obtained. In view of the close inter-relation between the temperature and current distributions, the turbidity of the water might be an important factor in the dynamics of wind currents.

Except in the case of Sweetwater Lake, where the winds were quite low, the computed depths of the thermocline are somewhat less than half the recorded depths. Perhaps this discrepancy is due to the assumed relationship $A_o(U)$ (Fig. 5) for the eddy coefficient under *indifferent* equilibrium which is based largely on measurements made when the density stratification was stable.

It has been demonstrated: that the sharp transition between the mixed layer and the thermocline can be accounted for theoretically by letting the eddy coefficients be functions of the stability and shear;

that the theory gives a thermocline depth which, although too shallow, is of the right order of magnitude; that this depth depends on wind speed, latitude, heat flux, and the *T-S* correlation, in the order stated, and that the variation with each of these factors occurs in a reasonable direction.

ACKNOWLEDGMENTS

We are indebted to Messrs. W. C. Jacobs, E. C. LaFond and B. E. Holtsmark for permission to use some of their unpublished data. Messrs. R. L. Stoker and E. Janssen of the Differential Analyzer Service, Department of Engineering, University of California at Los Angeles, have been helpful beyond what was involved in setting up the problem on the analyzer. Dr. L. Lek carried out some of the numerical computations, but unfortunately circumstances did not permit him to take a more active part and to join us in submitting this paper in appreciation to Dr. H. U. Sverdrup for having taught us oceanography.

REFERENCES

Brunt, David

1939. Physical and dynamic meteorology. Cambridge Univ. Press. 2nd ed. 428 pp.

Ekman, V. W.

1905. On the influence of the earth's rotation on ocean-currents. Ark. Mat. Astr. Fys., *2* (11): 1–52.

Helland-Hansen, Bjørn

1930. Physical oceanography and meteorology. Rep. Sars N. Atlantic Deep Sea Exped., 1910, *1:* 1–217.

Jacobsen, J. P.

1913. Beitrag zur Hydrographie der dänischen Gewässer. Medd. Komm. Havundersøg. Kbh., (Hydr.) *2:* 94 pp.

Rossby, C.-G. and R. B. Montgomery

1935. The layer of frictional influence in wind and ocean currents. Pap. phys. Oceanogr. Meteorol., *3* (3): 1–101.

Sverdrup, H. U.

1936. The eddy conductivity of the air over a smooth snow field. Geofysiske Publ., *11* (7): 1–69.

Sverdrup, H. U., M. W. Johnson and R. H. Fleming

1942. The oceans. Prentice-Hall, Inc., New York. 1087 pp.

Taylor, G. I.

1931. Internal waves and turbulence in a fluid of variable density. Rapp. Cons. Explor. Mer, *76:* 35–43.

STREAMLINES FROM A DISCRETE VECTOR FIELD: WITH APPLICATION TO OCEAN CURRENTS[1]

BY

DONALD W. PRITCHARD[2]

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

The problem of presenting a scalar field on the surface of the earth or sea can be solved by drawing the contour lines or isopleths of the field. If the values of the scalar are measured at points of a grid, it is necessary to interpolate between the points in constructing the contours. The resulting curves are therefore somewhat subjective; however, if the data are reasonably accurate and closely spaced, experienced workers can reduce this subjectiveness to a minimum.

The cartographic problem of representing a vector field, such as of surface currents or of winds, can be solved by drawing the streamlines. The art of interpolating between grid points and of drawing the streamlines is considerably more difficult than that of constructing the isopleths of a scalar field. The subjective element is stronger, especially when the data are widely spread and only moderately accurate. If the streamline problem could be reduced to one of contouring a scalar field, then much of the subjective nature of the problem would be removed. Also inherent in this problem is the necessity of developing an objective method of smoothing the data to reduce the effect of random observational errors.

In this paper a method is derived whereby the problem of drawing streamlines is reduced to one of contouring a scalar distribution; this method applies to the special case of an ocean without horizontal divergence. An indication of the method of attacking the more general case is given towards the end of the paper. The application of the special solution to the real ocean is discussed, and an example is given. The scalar quantity to be contoured in the special case of horizontal nondivergent flow is called the stream function. In contouring the stream function one has an advantage not encountered in contouring the usual scalar field, for the isopleths, or streamlines, should be parallel to the velocity vectors; and hence, besides using the numerical values of the stream function, certain details of the streamlines may be brought out by using the vectors themselves as directional guides. The contours of the stream function, however, give the general flow pattern which is frequently lost when only the directional qualities of the vectors are used directly.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 384. This work represents results of research carried out for the Hydrographic Office, the Office of Naval Research, and the Bureau of Ships of the United States Navy Department under contract with the University of California.

[2] Now with the Oceanographic Studies Branch, United States Navy Electronics Laboratory, San Diego, California.

STREAMLINES FOR A TWO-DIMENSIONAL NONDIVERGENT VELOCITY FIELD

In this special case the streamlines are contours of the so-called stream function. In the following discussion, consider a velocity vector field, with the vectors given at the points of a grid. The x and y components of the velocity are $u(x, y)$ and $v(x, y)$ respectively. With this notation, it can be shown (see Sokolnikoff, *et al.*, 1941: 221) that the stream function $\Psi(x, y)$ is related to the velocity field by

$$(1) \qquad \frac{\partial \Psi}{\partial x} = -v, \qquad \frac{\partial \Psi}{\partial y} = +u.$$

Solving these equations for the stream function,

$$(2) \qquad \Psi(x, y) = \int_{a,b}^{x,y} (-v dx + u dy).$$

The problem of finding the scalar function, the contours of which are streamlines, is thus reduced to the problem of numerical integration of equation (2). Consider the grid of velocity vectors described above to extend to n in the x-direction and to m in the y-direction. For every point in the field defined by an integer combination of the coordinates x and y, there exists a value of u and of v. Let $u_{i,j}$ and $v_{i,j}$ be the values of the components found at the intersection of the i^{th} column and the j^{th} row. For the purposes of the numerical integration, the following sums are introduced:

$$(3) \quad (a) \quad (\psi_y)_{l,\, j+1/2} = \sum_{j=0}^{k} u_{l,\, j} \Delta y \qquad \text{(For each } l,\ k \text{ takes on successive values from 1 to } m\text{)}$$

$$(b) \quad (\psi_x)_{i+1/2,\, k} = -\sum_{i=0}^{l} v_{i,\, k} \Delta x \qquad \text{(For each } k,\ l \text{ takes on successive values from 1 to } n\text{).}$$

Equation (3a) signifies the summation of u along a particular (l) value of x. For each particular value of l, k takes on successive values from 1 to m. Hence, for each value of l there are m equations of the form (3a), each one giving a value of ψ_y. There are n columns along which these partial summations are performed; that is, l takes on successive values from 1 to n. Hence, there are as many values of ψ_y as there are grid points. These values are assumed to apply along the grid lines parallel to the y-axis, and at points midway between the grid points. Similarly, equation (3b) signifies the summation of v along a particular (k) value of y, and for each of these values of k, l takes on successive values from 1 to n. The resulting mn values of ψ_x are assumed to apply along the grid lines parallel to the x-axis, and at points midway between the grid points. ψ_x equals zero along the line $x = ½$, and ψ_y equals zero along the line $y = ½$.

The discrete values of ψ_x and ψ_y thus found may be thought of as representing numerical values of a series of continuous functions, one such function of ψ_y applying for each of the columns of the grid, and one such function of ψ_x applying for each of the rows of the grid. These functions represent a sort of "line" stream function, applicable only for the particular velocity component summed, and for the line along which the summation is performed. We have at this point two fields of "line" stream functions; the numerical differentiation of one of these fields in the y-direction yields the x component of the velocity, and the numerical differentiation of the other field in the x-direction yields the y component of the velocity.

It is desired, however, to obtain *one* set of numbers in the field which represents one stream function, the differentiation of which in any particular direction yields the velocity component normal to that direction. Assume for the moment that the vector field being investigated is a true nondivergent field. In this case any particular row of values of ψ_x or column of values of ψ_y differs from the true values of the stream function along that particular row or column only by a constant of integration. This constant is different for each row and column. It is thus necessary to find the set of constants to be added to the rows of ψ_x's and to the columns of ψ_y's in order to make the two fields coincide.

Before proceeding with the finding of these constants it is necessary to remember that the numerical values of ψ_x and ψ_y are not given at the same points in the field, for ψ_x is given along integer values of y but at half-integer points of x, while ψ_y is given along integer values of x but at half-integer values of y. The values of both "line" stream functions for each integer combination of the co-ordinates are therefore found by interpolation.

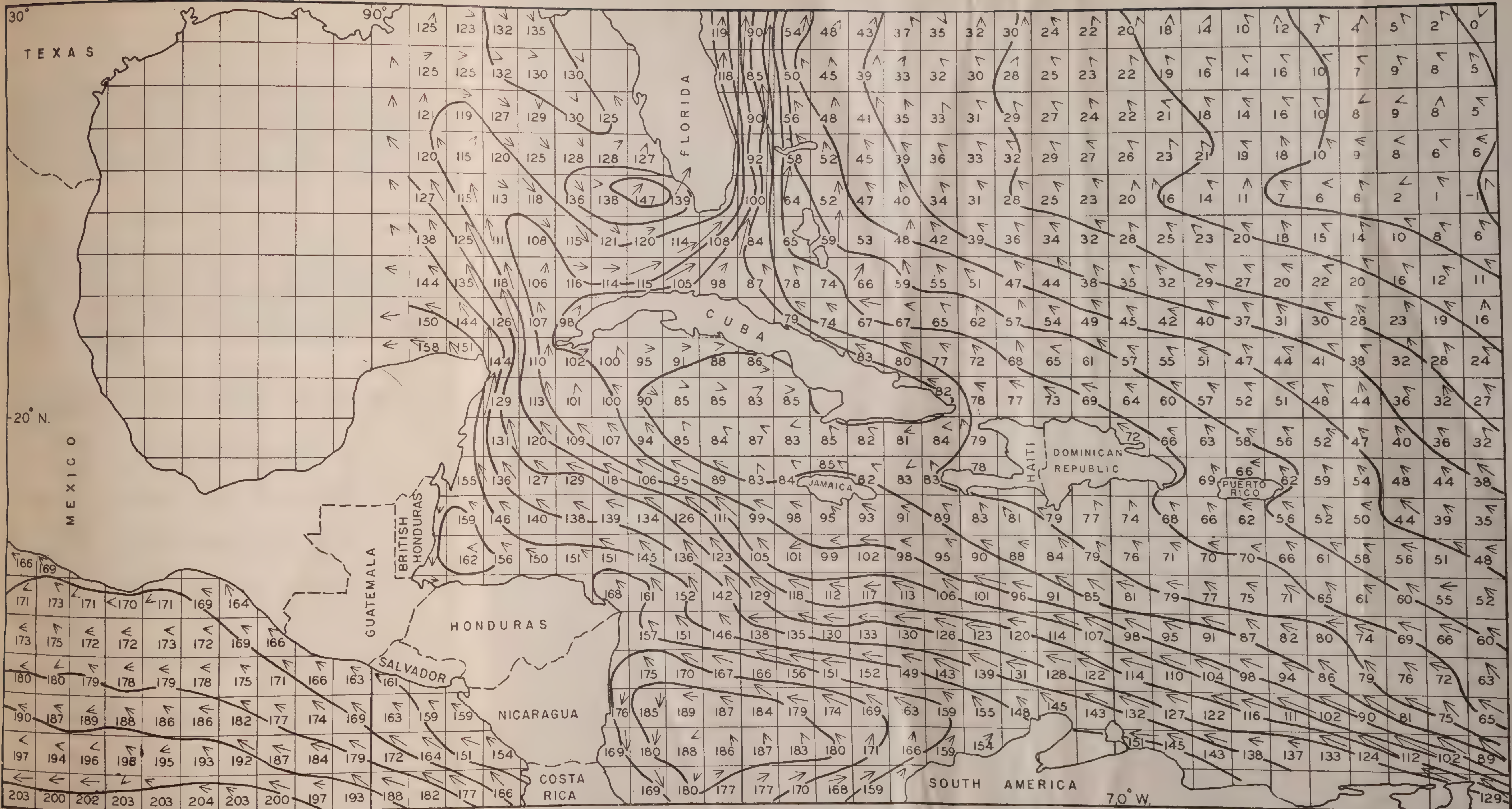

30°
90°
TEXAS
FLORIDA
CUBA
MEXICO
20° N.
BRITISH HONDURAS
GUATEMALA
HONDURAS
SALVADOR
NICARAGUA
COSTA RICA
JAMAICA
HAITI
DOMINICAN REPUBLIC
PUERTO RICO
SOUTH AMERICA
70° W.

FIGURE 1. Numerical values of the stream function, based on average drift per 24 hours, are shown in the center of each square. Solid lines are the streamlines. The vectors in each square represent the average observed drift per 48 hours.

At every integer combination of the co-ordinates x and y there are now two numbers, representing the numerical values of the two line stream functions at these points. The symbols ψ_x and ψ_y represent these numerical values. To obtain from these two sets of numbers one set of numbers, having the numerical values of the complete stream function for the field, a constant a_i is added to each row of ψ_x's and a constant b_j is added to each column of ψ_y's, such that

$$\psi_{x_{i,j}} + a_i = \psi_{y_{i,j}} + b_j \quad . \tag{4}$$

The value of a_i may differ from row to row and b_j may differ from column to column.

The condition represented by equation (4) can only be satisfied if the field is an ideal nondivergent field. In actual practice random errors in observation, even of a nominal nondivergent field, make it impossible to satisfy the equation (4) exactly. Instead, it is necessary to find statistically the set of constants which most nearly satisfies equation (4). This may be done by imposing the condition that the square of the difference

$$(\psi_{x_{i,j}} + a_i) - (\psi_{y_{i,j}} + b_j)$$

be minimized, both along rows and along columns. This condition yields the four equations:

(5) (a)

$$\text{(i)} \quad \frac{\partial}{\partial a_i} \sum_{j=1}^{n} \left\{ (\psi_{x_{i,j}} + a_i) - (\psi_{y_{i,j}} + b_j) \right\}^2 = 0$$

$$\text{(ii)} \quad \frac{\partial}{\partial b_j} \sum_{j=1}^{n} \left\{ (\psi_{x_{i,j}} + a_i) - (\psi_{y_{i,j}} + b_j) \right\}^2 = 0$$

(b)

$$\text{(i)} \quad \frac{\partial}{\partial a_i} \sum_{i=1}^{m} \left\{ (\psi_{x_{i,j}} + a_i) - (\psi_{y_{i,j}} + b_j) \right\}^2 = 0$$

$$\text{(ii)} \quad \frac{\partial}{\partial b_j} \sum_{i=1}^{m} \left\{ (\psi_{x_{i,j}} + a_i) - (\psi_{y_{i,j}} + b_j) \right\}^2 = 0$$

which reduce to

(6)

$$\text{(a)} \quad \sum_{j=1}^{n} \psi_{x_{i,j}} + n a_i = \sum_{j=1}^{n} \psi_{y_{i,j}} + \sum_{j=1}^{n} b_j$$

$$\text{(b)} \quad \sum_{i=1}^{m} \psi_{y_{i,j}} + m b_j = \sum_{i=1}^{m} \psi_{x_{i,j}} + \sum_{i=1}^{m} a_i \, .$$

Since the stream function is contained as a differential in the field it may be changed by a constant amount without affecting the differential properties which define the field. Therefore, one of the constants, either for a row or a column, may be given an arbitrary value. Here let $a_1 = 0$; *i.e.*, for $x = 1$, $a = 0$. Then from (6a), at $i = 1$,

$$\sum_{j=1}^{n} b_j = \sum_{j=1}^{n} \psi_{x_{1,j}} - \sum_{j=1}^{n} \psi_{y_{1,j}} \tag{7}$$

and (6a) becomes

$$a_i = \frac{1}{n}\left\{\sum_{j=1}^{n} \psi_{y_{i,j}} - \sum_{j=1}^{n} \psi_{x_{i,j}} + \sum_{j=1}^{n} \psi_{x_{l,j}} - \sum_{j=1}^{n} \psi_{y_{l,j}}\right\}. \tag{8}$$

Once all of the a_i's are found from (8), values of b_j may be computed for each column from

$$b_j = \frac{1}{m}\left\{\sum_{i=1}^{m} \psi_{x_{i,j}} - \sum_{i=1}^{m} \psi_{y_{i,j}} + \sum_{i=1}^{m} a_i\right\}, \tag{9}$$

which follows from (6b).

If the horizontal velocity field under consideration were ideally nondivergent and free from observational error, then only one set of constants would have to be found, since either of the two relations,

$$(10) \qquad \begin{aligned} &\text{(a)}\ \ \Psi\,(x, y)_{i,j} = \psi_{x_{i,j}} + a_i \\ &\text{or} \\ &\text{(b)}\ \ \Psi\,(x, y)'_{i,j} = \psi_{y_{i,j}} + b_j \end{aligned}$$

should define the same scalar field and represent the actual values of the stream function. In practice, however, it will be found that these two functions are not exactly the same, and the amount by which they differ is indicative of how much the actual field departs from the assumed nondivergent field. If the function as defined by (10a) is partially differentiated in both the x- and y-directions, it would yield both the observed values of the velocity component $v(x, y)$, and the best fit to the observed field of the component $u(x, y)$ that would result in a nondivergent field. Likewise, if the function as defined by (10b) is partially differentiated with respect to x and y, then the observed values of the component $u(x, y)$ will be obtained, as well as a set of values for the component $v(x, y)$, which is the best fit to the observed set that would result in a nondivergent field. It is convenient to average these two values for Ψ and thus obtain one set of numbers at each point in the grid representing the numerical value of the "stream function" for the field. The stream function defined in this manner is only one type of "best fit" to the observed velocity field.

If, for some reason, it is felt that one component of the velocity is likely to be more representative than the other, then a weighted average of the two functions defined by equation (10) might be formed.

STREAMLINES FOR A GENERAL TWO-DIMENSIONAL DIVERGENT VELOCITY FIELD

It is theoretically possible to express the streamlines of any two-dimensional velocity field by a mathematical equation of the form

$$\Psi(x, y) = \text{constant}, \tag{11}$$

the various lines being distinguished by different values of the constant. The existence of this stream integral, Ψ, makes it theoretically possible to reduce the streamline problem to one of contouring a scalar field.

This cannot be accomplished, however, by direct measurements. The stream integral must be calculated from the velocity components. In order to do this, an integrating factor $\lambda(x, y)$ must be found for the equations

$$\frac{\partial \Psi}{\partial x} = -\lambda v, \qquad \frac{\partial \Psi}{\partial y} = \lambda u. \tag{12}$$

Such an integrating factor always exists, and can be found by solving the differential equation

$$\frac{\partial}{\partial x}(\lambda u) + \frac{\partial}{\partial y}(\lambda v) = 0. \tag{13}$$

Since u and v are given only at discrete grid points, (13) must be solved numerically for λ. The values for λ are then inserted in (12), and these equations may be solved numerically for Ψ, the stream integral.

Inseparable from the solution for λ and Ψ is the problem of smoothing the data. The formal solution to this general problem has been indicated by Dr. Carl Eckart (personal communication). However, it does not appear likely that the equations can be solved in practice without the aid of modern electronic computing machines, and this is economically feasible only for large amounts of data.

APPLICATION TO SHIP'S SURFACE CURRENT OBSERVATIONS

Certain semiobjective aids in the construction of streamlines have been suggested (for example, see Bjerknes, 1911). In seeking a procedure whereby subprofessional personnel could prepare reproducible streamline charts, however, these methods do not suffice. The general solution to the stream integral problem seems at present too difficult to place on an economical computation basis. The method developed for a nondivergent velocity field may be applied, however, to those parts of the ocean where the divergence term is small and of a random nature.

It is the author's opinion that over certain large areas of the ocean the vertical shear of the vertical velocity component is a small and random term compared to the effects of the horizontal velocity field. Hence, in constructing streamlines for average velocity data, such as the monthly average surface currents for one-degree quadrangles prepared by the Hydrographic Office, the horizontal field may be assumed to be on the average a nondivergent field. This assumption would obviously fail in regions of known, large scale convergences or divergences.

Equations (3), (8), (9), and (10) have been employed in constructing the streamlines for two areas, one a 20 by 30 degree area in the Northwest Pacific, the other a 20 by 40 degree area including the Caribbean, the Gulf of Mexico, and part of the Saragasso Sea. This second case is given here as an example.

Hydrographic Current Chart Misc. No. 10,690-8 was used to obtain the basic data. It contains a velocity vector for every degree square. The grid of our model is in this case defined by the ½ degree lines of latitude and longitude. The Δy and Δx in equation (3) are each equal to one (degree). Fig. 1 is a chart of the area, and the vectors shown are the average of ships's surface current observations for the month of August.

The method of obtaining values of the stream function for contouring, as presented above, was tried for the entire area at once, with fair success in the Caribbean and the Florida Straits. The spread of the two functions defined by (10) was, however, quite large in the Gulf of Mexico, in the area off the west coast of Mexico, and in the Saragasso Sea region. The region was then broken up into five overlapping subregions, each subregion encompassing as little land as possible. The Gulf of Mexico was not included in this determination. Equations (3), (8), (9) and (10) were employed for each of the subregions. The areas were then fitted together, making use of the overlapping portions. The spread between the two functions defined by (10) was in almost all cases very small, and the subregions fitted together with but little break in continuity. The numbers in the center of each degree square in Fig. 1 represent the average of the two values of the stream function as given by (10a) and (10b). The streamlines, representing the contours of the average value of the stream function, are also shown in Fig. 1.

This application of the method of constructing streamlines indicates that it is best to subdivide fairly large areas into overlapping regions about 10 or 15 degrees square. Little additional labor is involved in breaking up the area, and the checking of the computations is made easier. The method appears to present a procedure whereby subprofessional personnel may prepare reproducible streamline charts of large ocean areas with but little professional supervision.

ACKNOWLEDGMENTS

Thanks are extended to Mrs. Alfreda Nixon and Miss Viola Bush for their assistance in carrying out the computations involved in testing the method. Dr. H. U. Sverdrup and Dr. Walter Munk gave considerable encouragement towards seeking a solution to this problem. I especially appreciate the time given by Dr. Carl Eckart in clarifying the over-all problem and in pointing out the relation of this special solution to the general case.

LITERATURE CITED

Bjerknes, V.
1911. Dynamic meteorology and hydrography. Pt. II. Kinematics. Publ. Carneg. Instn., No. 88: 1–175.
Sokolnikoff, I. S. and E. S. Sokolnikoff
1941. Higher mathematics for Engineers and Physicists. McGraw-Hill Book Co., Inc., New York. 587 pp.

A MODEL OF THE VERTICAL STRUCTURE OF MASS IN EQUATORIAL WIND-DRIVEN CURRENTS OF A BAROCLINIC OCEAN[1]

By

ROBERT O. REID

Scripps Institution of Oceanography
University of California
La Jolla, California

A model for the vertical distribution of density, which approximates observed conditions in the equatorial Pacific Ocean, is presented. This consists of a homogeneous layer near the surface overlying a baroclinic medium wherein the density varies exponentially in such a manner as to approach a constant value with increasing depth. It is shown that the depth of the upper homogeneous layer can be considered as a function only of the wind stress which is exerted upon the sea surface. The relationship is developed from Sverdrup's (1947) theory of wind-driven currents in a baroclinic ocean, where it has been demonstrated that the mean *horizontal* distribution of mass of the upper 1,000 meters is determined by the wind stress. The mean horizontal distribution of mass is, in turn, directly related to the depth of the homogeneous layer. Consequently, it is possible to describe the *three-dimensional* distribution of density and pressure, provided that the wind stress is known.

The distribution of density in a vertical section is constructed for the case of an east-west wind stress whose magnitude varies sinusoidally with latitude. The computed model is compared with a similar cross section of density obtained from oceanographic observations in the eastern Pacific.

THEORY

Review. In a baroclinic ocean the horizontal pressure gradient and velocity generally vanish at great depth. If d (a constant) is the selected depth of no motion and $p = p(xyz)$ is the pressure, then the

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 385. This work represents results of research carried out for the Hydrographic Office, the Office of Naval Research, and the Bureau of Ships of the Navy Department under contract with the University of California.

vertical integral of pressure $P = P(xy)$, which represents the mean horizontal distribution of mass, is defined by

$$\frac{\partial P}{\partial x} \equiv \int_{\text{Surface}}^{d} \frac{\partial p}{\partial x} \cdot dz \,, \quad \frac{\partial P}{\partial y} \equiv \int_{\text{Surface}}^{d} \frac{\partial p}{\partial y} \cdot dz \,, \tag{1}$$

where z is the vertical co-ordinate measured downward from a reference near the surface. For unaccelerated motion in a baroclinic ocean, free of lateral friction and transient changes of density, it can be shown on the basis of Sverdrup's (1947) theory that the quantity $P(xy)$ is given by the expression

$$P = (x - x_o) \left[\bar{\tau}_x - \frac{\partial \bar{\tau}_x}{\partial y} R \cdot \tan \phi \right] + \int_{y_1}^{y} \tau_y \, (y') \, dy' \,, \tag{2}$$

where $\bar{\tau}_x = \bar{\tau}_x(xy)$ and $\tau_y = \tau_y(y)$ are the x and y components of wind stress, respectively, in a rectilinear system where x and y are the horizontal co-ordinates positive to the east and north, respectively, with the origin on the equator. The term ϕ is the latitude, positive north of the equator, and R is the mean radius of the earth. The zero subscript refers to the continental boundary, and the *bar* superscript indicates an average value over the range x to x_o. The limit y_1 represents an arbitrary limit. Expression (2) is valid provided (a) that a constant depth of no motion, d, does exist, (b) that a north-south vertical boundary is present, and (c) that τ_y is independent of x.

It can be shown that (2) is a very good approximation for the eastern equatorial Pacific, even where x_o—the co-ordinate of the American continental boundary—varies with latitude.

The Model. The selected density structure is essentially a generalization of the simple two layer system treated by Defant (1936), where the upper homogeneous layer is retained. The structure is as follows: there is [1] an upper homogeneous layer of depth $h = h(xy)$ and density ρ_o; beneath this is [2] a baroclinic transitional layer extending from h to depth d, below which the water is at rest and of constant density ρ_d. An exponential transition will be assumed, such that the density $\rho(xyz)$ approaches ρ_d asymptotically with increasing depth, in good approximation of the model. One possible form is

$$\rho \, (xyz) \equiv \rho_2 = \rho_d - \Delta\rho e^{1-z/h} \;, \; z > h \,, \tag{3}$$

where $\Delta\rho = \rho_d - \rho_o$. The model is illustrated in Fig. 1, in which it is compared with actual distributions of ρ with depth, for various values of h. From (3) it is possible to solve for the P-function in terms of h.

To a very close degree of approximation the pressure is given by the hydrostatic equation,

$$p = g \int_{-\zeta}^{z} \rho \, (xyz) \, dz \, , \tag{4}$$

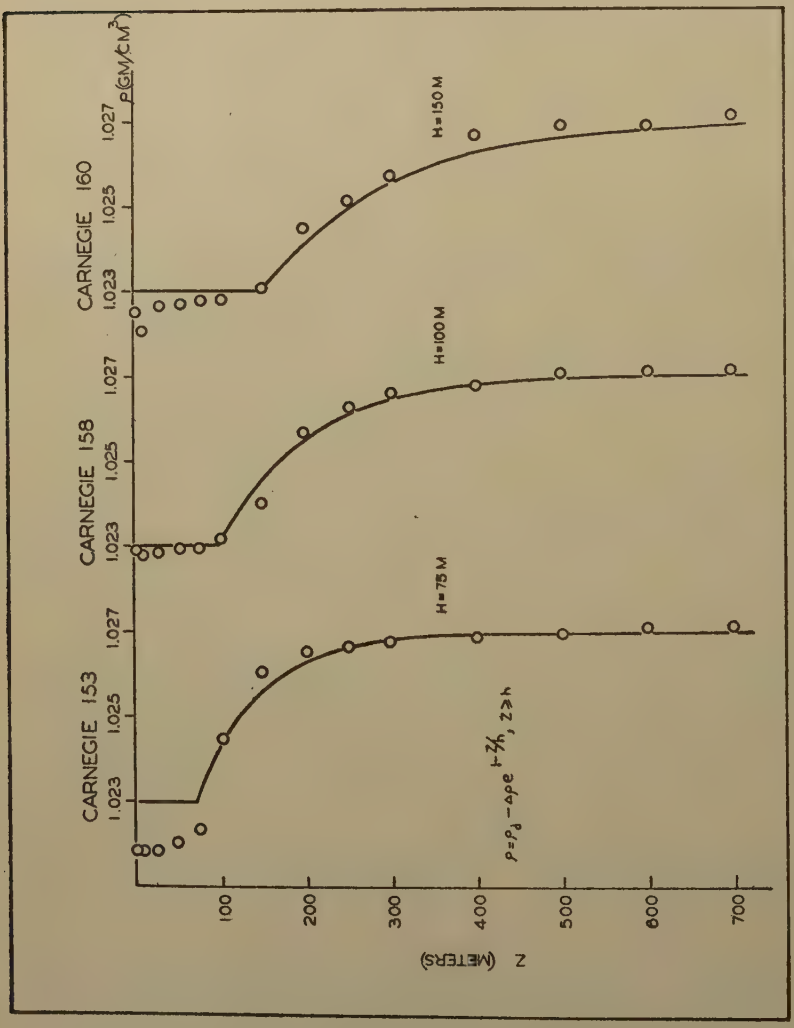


Figure 1. Comparison of observed densities (neglecting compressional effect) for three Carnegie stations in the Pacific with the computed vertical distribution of density for various values of h.

where $\zeta = \zeta(xy)$ is the *elevation* of the surface, and g is the acceleration of gravity.

It will be convenient to deal with the space derivatives of pressure rather than with the pressure itself. The pressure gradient in the x-direction is given by

$$\left(\frac{\partial p}{\partial x}\right)_1 = g\frac{\partial}{\partial x}\int_{-\zeta}^{z}\rho_o dz = g\rho_o\frac{\partial \zeta}{\partial x}, \quad -\zeta \leqslant z \leqslant h, \tag{5}$$

in the upper layer, and by

$$\left(\frac{\partial p}{\partial x}\right)_2 = g\left[\rho_o\left(\frac{\partial \zeta}{\partial x} + \frac{\partial h}{\partial x}\right) - \rho_d\frac{\partial h}{\partial x} + \Delta\rho\frac{\partial}{\partial x}\left(he^{1-z/h} - h\right)\right], \quad z \geqslant h, \tag{6}$$

in the transition layer. Similar expressions are valid for $\partial p/\partial y$.

At $z \geqslant d$, the pressure gradient vanishes, and if $d >> h$, (6) yields the approximation

$$\zeta = \frac{2\Delta\rho}{\rho_o}h, \tag{7}$$

provided that the reference level is selected properly. Both h and ζ must either be positive or vanish.

The derivative of integrated pressure is, from (1), (6) and (7)

$$\frac{\partial P}{\partial x} = \int_{-\zeta}^{h} g\rho_o\frac{\partial \zeta}{\partial x}\cdot dz + \int_{h}^{d} g\Delta\rho\frac{\partial}{\partial x}\left(he^{1-z/h}\right)\cdot dz\,.$$

This yields the approximation

$$\frac{\partial P}{\partial x} = \frac{5g\Delta\rho}{2}\frac{\partial h^2}{\partial x}.$$

A similar expression holds for $\partial P/\partial y$, and it follows

$$P = \frac{5}{2}g\Delta\rho\cdot h^2 + \text{constant.} \tag{8}$$

The Sinusoidal Wind Stress Distribution. The east-west component of wind stress in the equatorial Pacific has been shown to vary nearly sinusoidally with latitude (Reid, 1948). It will be assumed here for simplicity that

$$\bar{\tau}_x = \tau_x = \tau_o + b\sin\omega(\phi - \phi_o),\ \tau_y = 0, \tag{9}$$

where τ_o, b, ω, ϕ_o are constants. The function P and consequently h

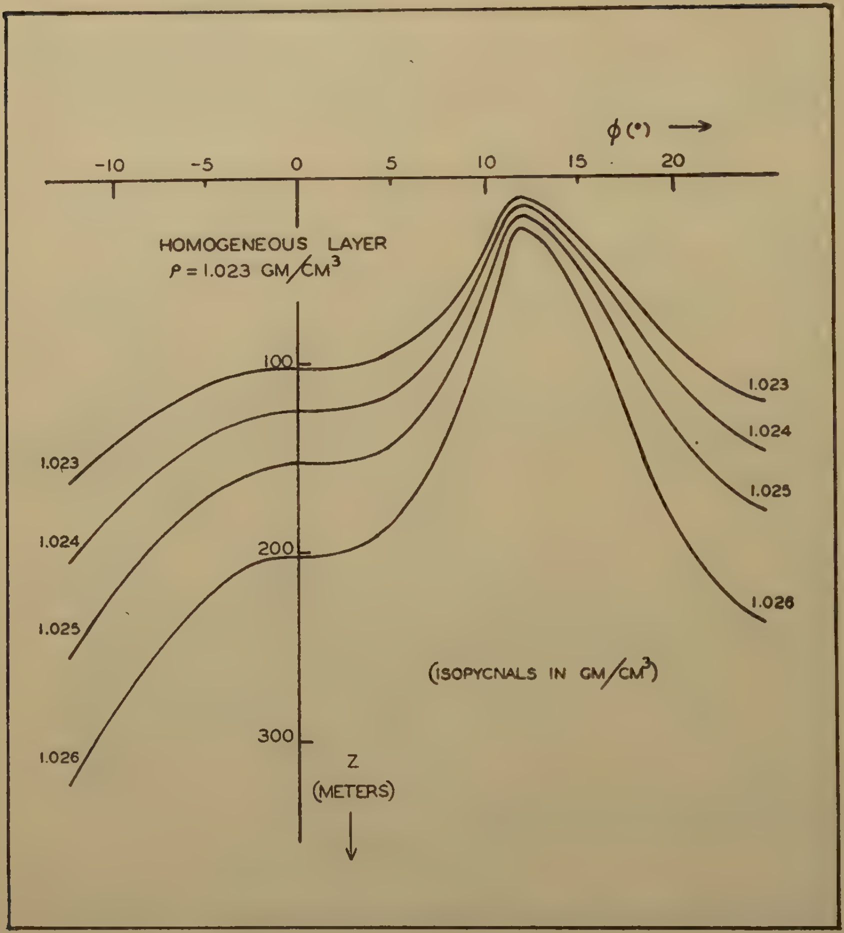


Figure 2. The computed distribution of density along the CARNEGIE section of the central equatorial Pacific, derived from the hypothetical exponential density model and the sinusoidal wind stress distribution.

can be evaluated in terms of the position co-ordinates. Using (2), (8), and (9) it follows

$$h^2 = \frac{2}{5g\Delta\rho}(x - x_o)\left[\tau_o + b\sin\omega(\phi - \phi_o) - \omega b\tan\phi \cdot \cos\omega(\phi - \phi_o)\right] - K, \qquad (10)$$

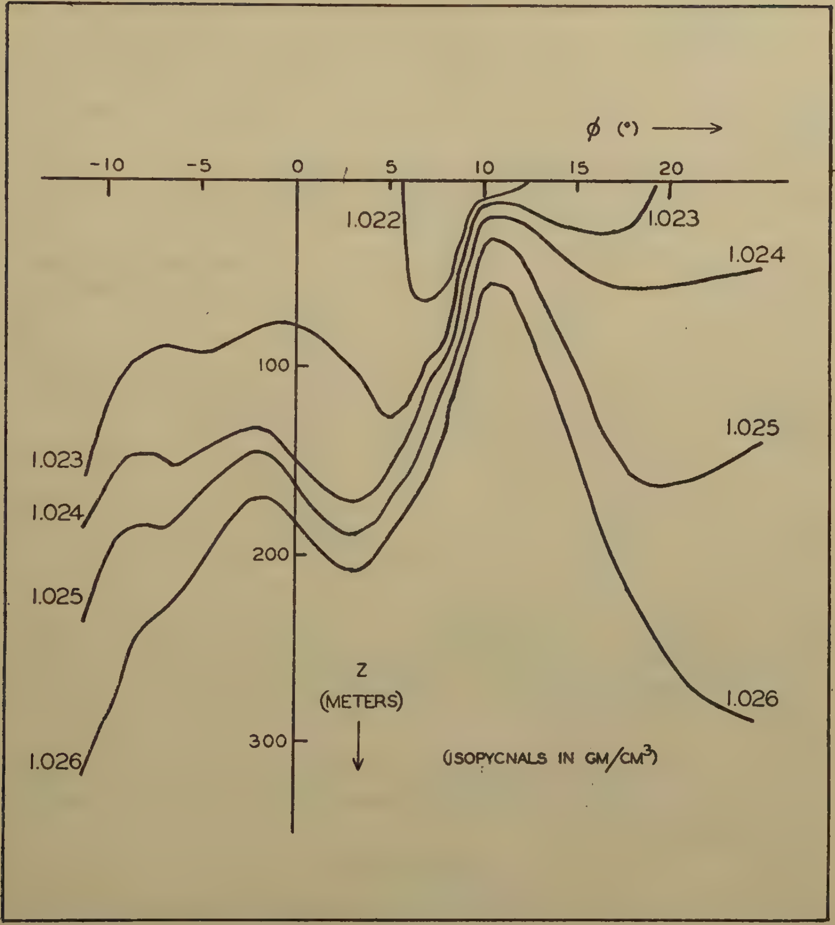


Figure 3. The actual distribution of density along the CARNEGIE section of the central equatorial Pacific, derived from oceanographic data.

and the surface elevation is

$$\zeta^2 = \frac{8\Delta\rho}{5g\rho_o^2}(x - x_o)\left[\tau_o + b\sin\omega(\phi - \phi_o) - \omega b\tan\phi \cdot \cos\omega(\phi - \phi_o)\right] - \left(\frac{2\Delta\rho}{\rho_o}\right)^2 K, \qquad (11)$$

where K is the constant of integration, to be selected such that h is always a real quantity.

It is evident that h^2 and ζ^2 *increase linearly* with distance from the boundary for the chosen distribution of density and wind stress.

The Pressure Distribution. The form of the pressure distribution is obtained readily from the hydrostatic equation and the model (3), giving

$$\left.\begin{aligned} p_1 &= g\,(\rho_o z + 2\Delta\rho h), \; -\zeta \leqslant z \leqslant h, \\ \text{and} \quad p_2 &= g\,(p_d z + \Delta\rho h e^{1-z/h}), \; z \geqslant h. \end{aligned}\right\} \tag{12}$$

The slope of the isosteres, $i\rho_2$, in the transition layer can be shown to be

$$i\rho_2 = \left(\frac{z}{h}\right) i_h\,, \quad z \geqslant h\,, \tag{13}$$

where i_h is the slope of the interface at $z = h$. It can be verified that the isobaric slopes, i_p, for the two layers are

$$\left.\begin{aligned} i_{p_1} &= -\frac{2\Delta\rho}{\rho_o}\, i_h = i_\zeta\,, \\ \text{and} \quad i_{p_2} &= -\frac{\Delta\rho}{\rho}\left(1 + \frac{z}{h}\right)e^{1-z/h}\, i_h \doteq \frac{1}{2}\left(1 + \frac{z}{h}\right)e^{1-z/h}\, i_{p_1}\,. \end{aligned}\right\} \tag{14}$$

The isobaric slope in the transition layer is opposite in sign to the slope of h for all $z \geqslant h$. The isosteric slope, on the other hand, has the same sign as that of h for $x \geqslant h$. This clearly demonstrates the baroclinicity of the model fluid, *i. e.*, the isosteres intersect the isobars at all $z \geqslant h$.

It is evident from (14) that the isobaric slope becomes vanishingly small as z approaches d, which is necessary to satisfy the condition of stationary water below d.

DISCUSSION

The observations utilized in the computations consist essentially of a line of 13 stations between Lat. 24° 57′ N, Long. 137° 44′ W and Lat. 10° 54′ S, Long. 161° 53′ W, occupied by the Carnegie between October 21 and November 4, 1929 (Fleming, *et al.*, 1945).

A value of K was selected as 3.8×10^8 so that the minimum value of h for the model would correspond to that observed at about Lat. 13° N. The depth of no motion was taken as 700 m, and the mean values of ρ_o and ρ_d were taken for the Pacific data as 1.0230 and 1.0274 gm/cm³, respectively.[2]

---

[2] The values represented here do not include the effect of compression, nor does the model. The effect is unimportant, since there is very little contribution to the gradients of p and P due to compression.

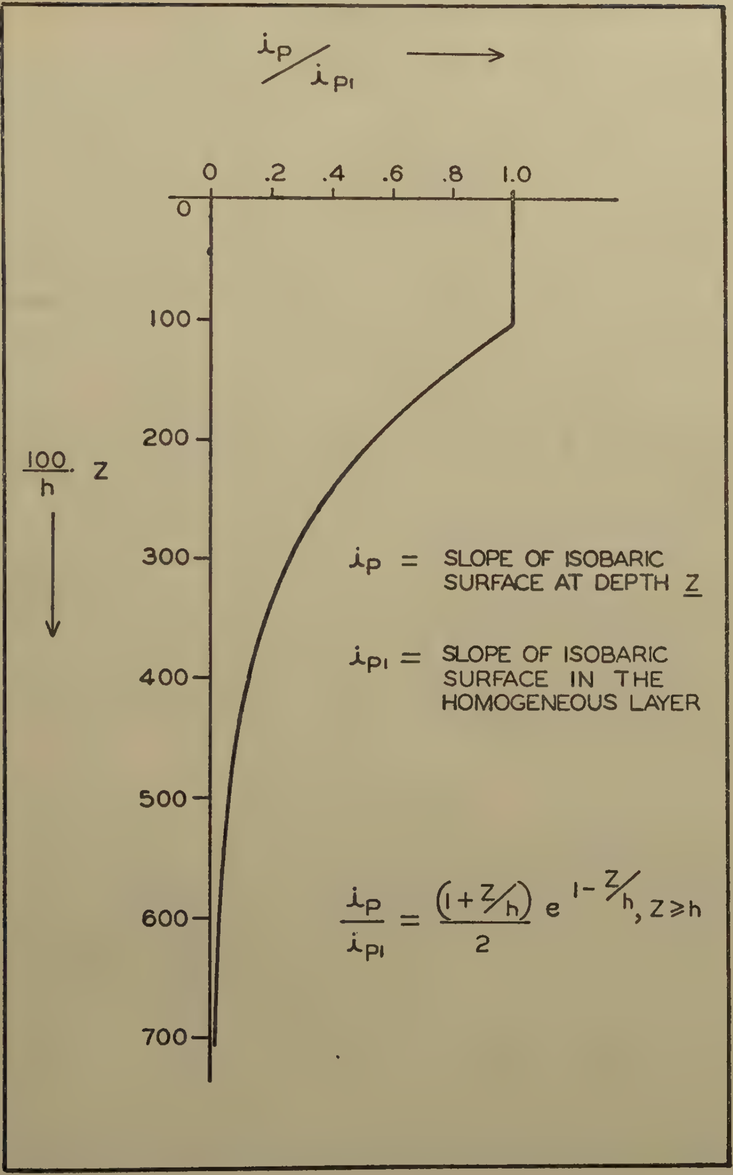


Figure 4. The dependence of isobaric slope on depth, for the exponential density distribution. The ratio of the slope to that in the homogeneous layer is plotted vs. depth in per cent of the homogeneous layer depth.

The values of the wind stress constants are:

$$\tau_o = -0.3 \text{ dyne/cm}^2, \; \omega = 13.85,$$
$${}_b = \;\; 0.3 \text{ dyne/cm}^2, \; \phi_o = 1°.$$

The hypothetical vertical density structure along the Carnegie section computed from expressions (3) and (10) and the above constants is shown in Fig. 2. The values of $(x - x_o)$ were obtained from the position of the Carnegie stations and the American continental boundary. The actual distribution of density, omitting effects of compression, is shown in Fig. 3 for comparison. A plot of the vertical variation of isobaric slope, obtained from (14), is shown in Fig. 4. There is no discontinuity in the slope at $z = h$, as in the case of the simple two layer system [where $\rho(z)$ is discontinuous].

SUMMARY AND CONCLUSIONS

The exponential transition of density below a light homogeneous layer is strikingly representative of the vertical distribution of density in the equatorial Pacific. From this model and the theory of wind-driven currents, the entire field of density and pressure is described. A cross section of the model density structure is compared with the actual distribution given by the Carnegie data. The principal characteristics of the observed structure of mass are accounted for in the model, except in the vicinity of the equator.

The conditions at the equator need further consideration, especially regarding the role of thermodynamic processes in determining the distribution of the P-function.

REFERENCES

Defant, Albert

1936. Schichtung und Zirkulation des atlantischen Ozeans. Die Troposphäre. Wiss. Ergebn. dtsch. Atlant. Exped. *Meteor*, 6 (1), Lief. *3:* 316–324.

Fleming, J. A., C. C. Ennis, H. U. Sverdrup, S. L. Seaton, and W. C. Hendrix

1945. Scientific results of cruise VII of the *Carnegie*, 1928–29. Oceanography I-B. Observations and results in physical oceanography. Graphical and tabular summaries. Publ. Carnegie Inst., *545:* 1–315.

Reid, R. O.

1948. The equatorial currents of the eastern Pacific as maintained by the stress of the wind. J. Mar. Res., *7* (2): 74–99.

Sverdrup, H. U.

1947. Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the eastern Pacific. Proc. nat. Acad. Sci., *33:* 318–326.

METHODS FOR ESTIMATING DYNAMIC SLOPES AND CURRENTS IN SHALLOW WATER[1]

BY

PIER GROEN

Section of Oceanography
Koninklijk Nederlands Meteorologisch Instituut

In computing relative currents in the sea from the mass distribution, a well known difficulty is encountered when a section extends into relatively shallow water; then a mass of land becomes part of the section. To compute the relative currents or the relative slope of any isobar of the section with respect to, say, the 500 db isobar, the ordinary rules cannot be applied for points below which the pressure at the bottom is less than 500 db.

First Method. To overcome this difficulty, Helland-Hansen (1934) proposed a method which replaces the area of the cross section occupied by the solid earth by a fictitious section of water within which the lines of equal specific volume are horizontal (or isobaric[2]) continuations of the corresponding isosteric curves in the water.

Second Method. Sverdrup, Johnson and Fleming (1942: 451) proposed a different method, which they describe as follows: "As another approximation, one can apply the equation $i_p = -\bar{\imath}_\delta(\delta_1 - \delta_2)$ (p. 412), *introducing the slope of the δ surfaces where they cut the bottom and taking the difference* $\delta_1 - \delta_2$ *along the bottom*" (italics mine). In the formula cited, i_p is the relative slope of any isobar $p = p_1$ with respect to a reference isobar $p = p_2$, δ_1 and δ_2 designate the specific volume anomaly δ at $p = p_1$ and at $p = p_2$ respectively, and $\bar{\imath}_\delta$ is the average inclination of the δ-lines between p_1 and p_2.

In a more general way we might describe the essence of this method as follows: In order to compute the relative slope of any isobar $p = p_1$ with respect to the reference isobar $p = p_2$ at a certain point P of the section, one uses the values of the specific volume α (or the anomaly δ) and the inclinations of the α-lines (or δ-lines[3]) along the vertical through P down to the bottom; beneath the bottom one projects horizontally onto this vertical the α-values (or the δ-values) and the inclinations of the α-lines (or δ-lines) occurring along the bottom-line in the cross section down to $p = p_2$.

---

[1] Although the author had no discussions with Professor Sverdrup on the present matter during a visit of three and a half months at the Scripps Institution, it is a great pleasure to express sincere thanks for the stimulating contacts and to contribute this small note in his honor.

[2] The difference between "horizontal" and "isobaric" is unimportant in this connection.

Now this procedure does not have the disadvantage of using a hypothetical physical picture of a rather arbitrary and somewhat unlikely sort (the barotropic, fictitious water-mass, which Helland-Hansen introduces). It has a disadvantage, however, in not corresponding to any physical picture at all, for it is impossible to replace the block of solid earth by a water mass such that the relative slopes of the isobars would everywhere be the same as the slopes computed by Sverdrup's method, *except* if the slopes of the isosteric lines should happen to be *zero* at all points of the bottom-line.

Indeed, in order to obtain a physical picture of this method of computation we would have to substitute for the block of solid earth a water mass which has along every vertical the horizontally projected values of the specific volume α *and* of the inclination of the α-lines occurring along the bottom; consequently, along any horizontal (or isobaric) line (any dotted line in Fig. 1) in this fictitious water mass, α as well as the inclination of the α-lines would have to be constant. Such a water mass could exist physically only for the special case where the α-lines in the actual water meet the sloping bottom horizontally.

Proposed Third Method. Fixing the inclination of the α-lines at every point of the fictitious water mass *alone* determines a whole α-field (if only the values of α along the bottom-line are given). Therefore, we may propose, as a third method, the following procedure:

The block of solid earth is replaced by a fictitious section of water, in which the density distribution is determined by putting the inclination of the α-lines along every horizontal (or isobaric) line equal to the value projected horizontally (or "isobarically") from the bottom-line. The α-lines are then extended from the actual water into the fictitious water by reference to these inclinations. The α-lines determined in this manner are shown as dashed lines in Fig. 1.

Instead of the specific volume we may also use its anomaly δ; the lines in Fig. 1 then represent δ-lines. If, for instance, we use the *approximate* formula cited above,

---

[3] Of course, the α-line and the δ-line have different slopes at any given point, which to use depends on the formula used.

$$i_p = -\bar{i}_\delta\,(\delta_1 - \delta_2),$$

then the difference of this method from Sverdrup's will be in the use of a different value of δ_2, namely a smaller one, if δ in the actual water increases away from the bottom along the reference level, and a larger one in the opposite case.

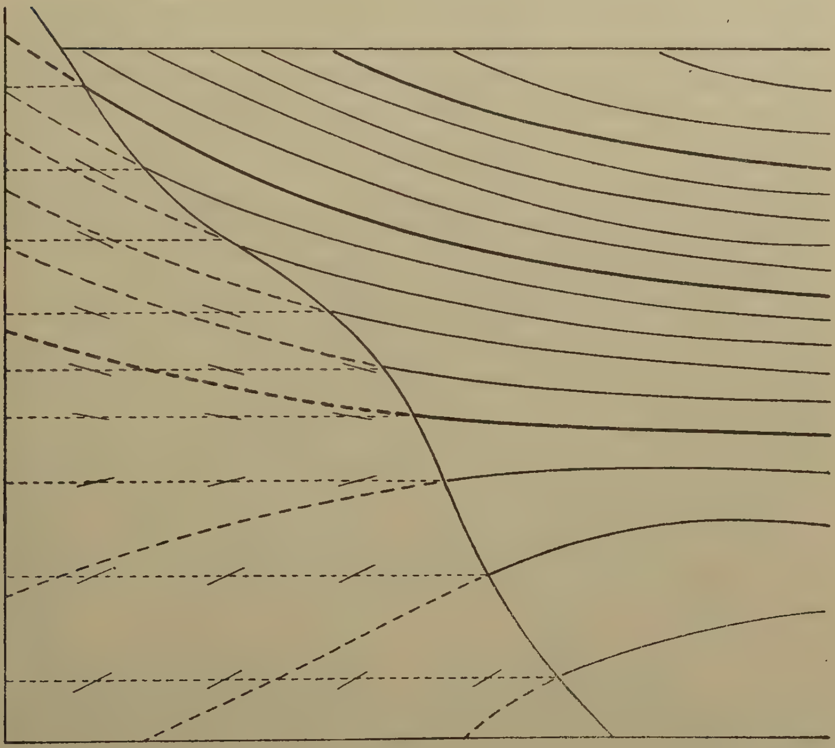


Figure 1. Vertical cross section extending into shallow water. The scales are arbitrary. The fully drawn curves are α-lines; their continuations to the left of the bottom-line (dashed curves) have been drawn in such a way that at each level they have the same slope (indicated by the straight little lines in the picture) as the α-line which meets the bottom-line at this level.

Discussion. Finding the value of the specific volume along a vertical in the fictitious water mass is not a matter in which high accuracy must be aimed at, for this third procedure does not, of course, pretend to give the *correct* answer.

As a matter of fact, no method of solving the problem can prove its correctness, since, strictly speaking, the problem is not exactly defined. Since the question is to find the relative slope of a certain isobar at a certain point with respect to a reference isobar that does not exist vertically below this point, the only thing we may do is to draw up a procedure which is not too complicated, and which, if possible, has some physical background that is not too unlikely.

Finally, we should not think that, when knowing the relative slopes or currents at two different points at the same level with respect to *one* reference isobaric surface, we also know the relative slopes or currents at those points with respect to each other, *unless* we know, from other sources, the relative current distribution within the reference surface.

REFERENCES

HELLAND-HANSEN, BJ.,
1934. The Sognefjord section. James Johnstone Memorial Volume. Liverpool, Univ. press: 257–274.
SVERDRUP, H. U., M. W. JOHNSON AND R. H. FLEMING
1942. The oceans. Prentice-Hall, Inc., New York. 1087 pp.

SOME RECENT DEVELOPMENTS IN THE STUDY OF THE GULF STREAM

By

C. O'D. ISELIN AND FREDERICK C. FUGLISTER

Woods Hole Oceanographic Institution[1]
Woods Hole, Massachusetts

INTRODUCTION

At Cape Hatteras the Gulf Stream leaves the Continental Slope, so that as far as the deep current is concerned its course is no longer confined. The current again approaches the Continental Shelf off Georges Bank and off the Grand Banks, but in neither area does it flow normally along the Continental Slope as a river pressing against its bank. The band of water which separates the Gulf Stream from the Continental Shelf in the sector between Cape Hatteras and the Grand Banks is usually called "slope water." The great body of water to the south and east of the Stream is known as the Sargasso Sea or Central Atlantic water. Although it is not a part of this paper to discuss the physical characteristics and the similarities or dissimilarities of these two masses of water, there is an important difference that must be mentioned. In the slope water, the main thermocline lies much nearer to the surface than it does in the Sargasso Sea. The depth of this thermocline does not gradually change from one body of water to the next but changes abruptly within a relatively short distance. The Gulf Stream occupies this transition belt where the midpoint of the thermocline drops from an average depth of 250 meters in the slope water to an average of 750 meters in the Sargasso Sea.

In the sector between Cape Hatteras and the Grand Banks, the Gulf Stream resembles a "river in the ocean" rather than a current pressed close against the margin of the ocean basin. For some purposes it is more convenient to think of the Gulf Stream as a transition belt or as a boundary phenomenon. The approximate mean position of this boundary, as taken from the U. S. Hydrographic Office chart No. 1412, is shown in Fig. 1. The dots in the figure show where the available subsurface temperature observations have indicated the abrupt change in the depth of the thermocline.

---

[1] Contribution No. 451 from the Woods Hole Oceanographic Institution.

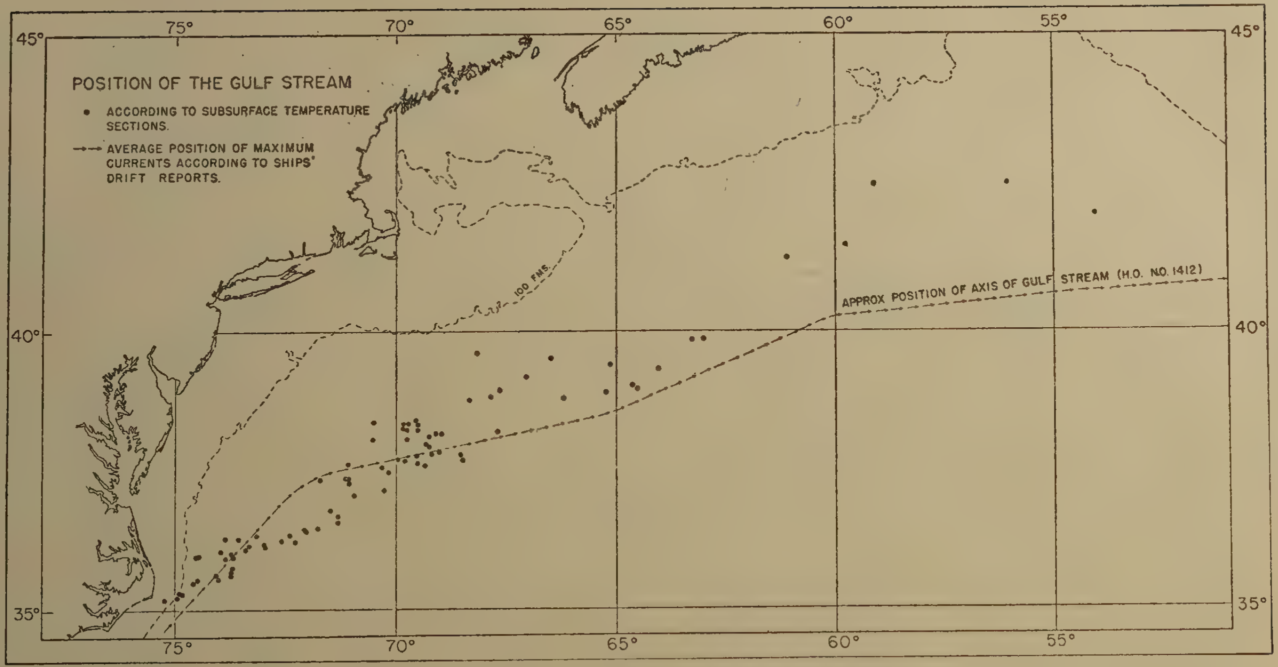

POSITION OF THE GULF STREAM
● ACCORDING TO SUBSURFACE TEMPERATURE SECTIONS.
—•— AVERAGE POSITION OF MAXIMUM CURRENTS ACCORDING TO SHIPS' DRIFT REPORTS.
APPROX POSITION OF AXIS OF GULF STREAM (H.O. NO. 1412)
100 FMS.
75°
70°
65°
60°
55°
45°
40°
35°

Figure 1. Comparison between the average path of the axis of the Gulf Stream as deduced from ship reports at the U. S. Hydrographic Office (arrows) and cross current temperature profiles (dots).

Our understanding of this part of the Gulf Stream system has come from three quite different sorts of studies, and only recently has a radically new attack on the problem been made. There is, first of all, a long series of studies in which average conditions in and near the Gulf Stream have been shown in various ways. The mean position of the axis of the Stream as given in Fig. 1 is an example. As more data became available, these plottings of the average physical characteristics of the current have been greatly improved, either by shortening the time intervals (Fuglister, 1947) or by combining several different types of information (Neumann and Schumacher, 1944). As will be shown below, although these observations were useful as a preliminary step in understanding the Gulf Stream, all such studies in the process of averaging the available data show the current as being considerably broader and weaker than it actually is at any given time. Of course, the defect has been generally understood by oceanographers, but how pronounced it could be has only recently been demonstrated.

A second general method, which was employed vigorously during the period 1930–1940, was the study of profiles of the Gulf Stream derived from lines of serial temperature and salinity observations. These Gulf Stream sections required only two or three days to complete and thus approached the synoptic ideal, but the interpolations, both vertical and horizontal, were relatively great. Especially when in and near the current, the observer experienced great difficulty in controlling the navigation of the research vessel, and the placing of the stations in relation to the current was always in doubt. This method when used under the most favorable circumstances provided considerable evidence (Iselin, 1940) that in cross section the current behaves just as would be predicted from the circulation theorem. However, because of the relatively wide spacing of the observations, calculations of the velocity could be regarded only as minimum values, and there remained even greater doubt as to the velocity at considerable depths below the surface. The total transport of the current could be computed from such observations with somewhat greater assurance. In these Gulf Stream profiles, marked distortions of the main thermocline layer on either side of the current were encountered. In the absence of reliable surface current observations such undulations of the isotherms could only be tentatively ascribed to eddies. Furthermore, in each case only a single profile was obtained.

At the same time that these subsurface studies were being made, the surface temperature records from commercial ships making regular runs across the area were being obtained. From these data, Church (1937) concluded that the Gulf Stream wandered or meandered to an increasing degree as it moved eastward from Cape Hatteras. Although the surface temperature alone is not always an exact indication of the location of the swiftest current, this profitable line of study probably was not continued as long as it should have been.

A third approach to the problem of the Gulf Stream has stemmed from theoretical studies. The contribution of the basic ideas in the circulation theorem has already been mentioned. Rossby (1936) pointed out that the Gulf Stream, at least to some extent, must behave as a wake stream and that countercurrents and eddies should be found on both sides of it. More recently Stommel (1948) has contributed much to our understanding of why the great North Atlantic eddy, of which the Gulf Stream is a part, has been found to be so much swifter and narrower in the west than in other sectors. As will be shown below, the new observations are not inconsistent with these various theories, but as might be expected, they present us with a somewhat more complicated picture. It is hoped that in time the theory will be refined to meet the new aspects.

NEW INSTRUMENTS

During the war two new instruments were developed that allowed a different approach to the study of the Gulf Stream. These instruments were the bathythermograph, with which subsurface temperature data can be obtained while a ship is underway, and the loran, which makes it possible to obtain navigational fixes at any hour of the day or night.

With the bathythermograph, a nearly continuous temperature-depth profile can be obtained. Within a few minutes of the time of lowering this instrument, the observer has a clear picture of the temperature structure of the water through which the ship is passing. With this knowledge, he can maneuver the ship so as to remain close to any particular thermal structure that is to be studied. It is true that the instrument shows only the surface layer temperatures, to a depth of 900 feet with the latest model bathythermograph, but as will be shown, it is evident that even in this shallow layer the temperature structure reflects the current pattern.

The loran makes it possible to maneuver a ship in and out of a current over a period of days without becoming hopelessly involved in navigational problems. Unfortunately this does not hold for every portion of the Gulf Stream system. In many areas over the ocean, the loran coverage is poor or nonexistant, but it is fortunate that the coverage is extremely good over the area traversed by the Gulf Stream between Cape Hatteras and Georges Bank. Even in this region, a single loran fix should not be relied on, but if the instrument is in operation continuously and if fixes are obtained at frequent intervals, the results are far more reliable and useful than the best dead reckoning and celestial navigation.

Other instruments have been and are being developed which will be of value in ocean current studies. For example, there are new electrical means of recording surface salinity, and the multiple sea sampler gives promise of providing practical means of observing the subsurface distribution of salinity at frequent intervals. It is hoped that these devices will be used in future surveys of the Gulf Stream, but since they were not employed in these studies, they are mentioned only briefly.

RECENT INVESTIGATIONS

The first studies of the Gulf Stream made with a bathythermograph were of rather limited scope (Spilhaus, 1940). As the instrument was further developed, a number of sections were run across the Stream, and it became quite clear that the temperature structure in the surface layer (450 or 900 feet) indicated the presence of the current. Fig. 2 shows two such temperature profiles, one made in August 1945 and the other in December 1947. A study of these and other sections shows that the depth of the 65° F. isotherm can be used conveniently to indicate the position of the Gulf Stream as well as large eddies and countercurrents. In the Sargasso Sea, the temperature of the water at a depth of 900 feet is at all seasons of the year between 64° and 65° F. There is no evidence of the water at this depth ever becoming appreciably warmer than this, but on the other hand a strong eddy or current may cause it to become several degrees cooler. In the swiftest portion of the Gulf Stream, the 65° isotherm abruptly moves up to the surface. During the coldest season of the year, all slope water is colder than 65°, while at other times of the year this isotherm is generally less than 200 feet deep. Anticyclonic eddies in the slope water area will cause the 65° isotherm to deepen, but rarely will it go to as great a depth as it occupies in the Sargasso Sea.

In spite of the fact that these temperature profiles are comparatively easy to obtain, only a few complete cross current sections are yet available. Shallow temperature profiles show where a strong current has been encountered, but they alone do not permit any calculations of velocity. With the introduction of the loran, not only was the problem of navigation near a current much simplified, but a means of approximating the velocity of the current at the surface was supplied. The loran data confirmed what had long been suspected on theoretical grounds; namely, that the sharp drop of the isotherms in the surface layer occurs where the current is swiftest.

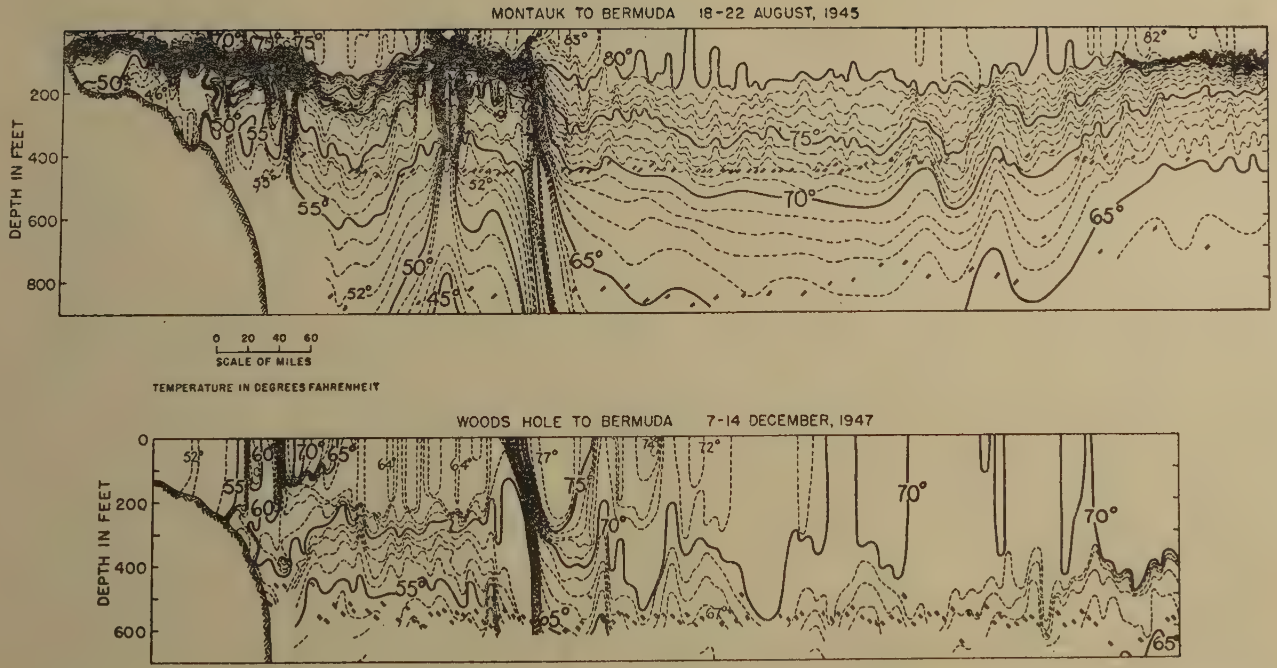


Figure 2. Bathythermograph profiles across the Gulf Stream off southern New England. The diagonal marks indicate the greatest depth achieved on each lowering. In the upper section the 450-foot instrument was used alternately with the 900-foot instrument.

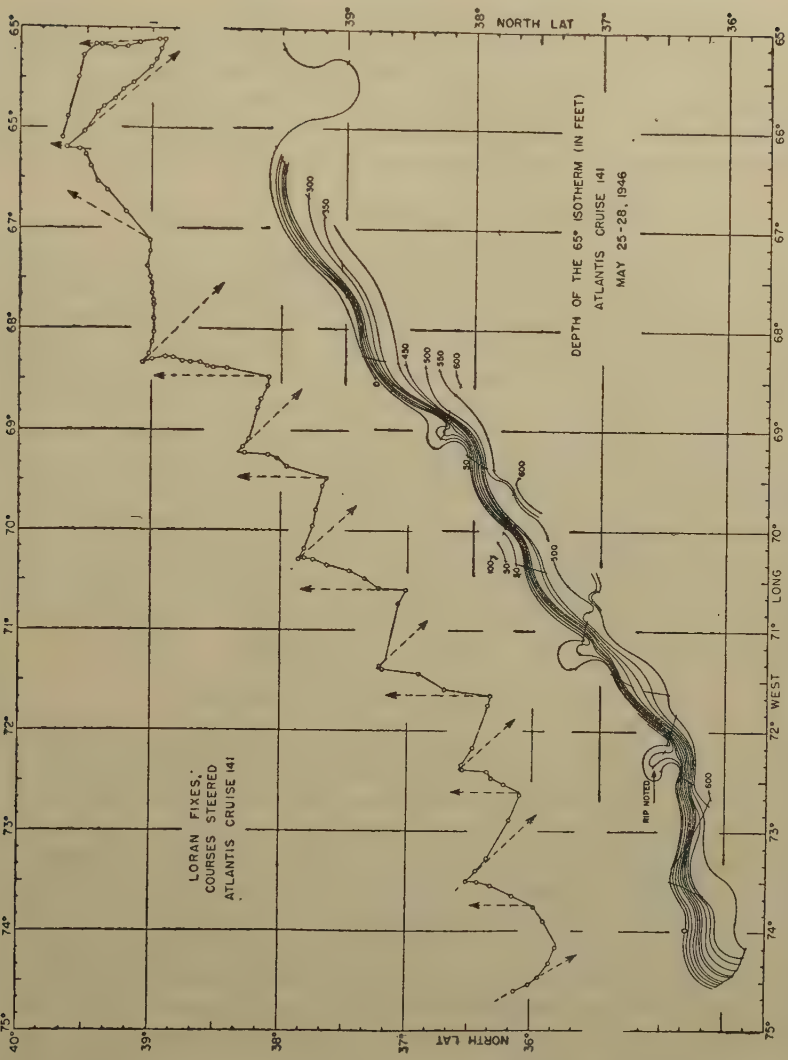

NORTH LAT
39°
38°
37°
36°
65°
66°
67°
68°
69°
70°
71°
72°
73°
74°
75°
40°
LONG
WEST
DEPTH OF THE 65° ISOTHERM (IN FEET)
ATLANTIS CRUISE 141
MAY 25-28, 1946
LORAN FIXES; COURSES STEERED ATLANTIS CRUISE 141
RIP NOTED
300
350
450
500
600
100

Figure 3. Survey of the northern edge of the Gulf Stream showing the departure between dead reckoning and observed positions (above) and the probable depth contours of the 65° F. isotherm (below).

In May 1946, the two new instruments were used aboard the research ketch ATLANTIS in a new type of survey of the Gulf Stream.[2] Instead of observing an individual section, a series of short sections was run, crossing and recrossing the Stream, the ship drifting with the current; in this manner a large area was covered and the course as well as the velocity of the Stream was determined. This series of sections is shown in Fig. 3. The upper portion of the diagram shows the loran fixes and the courses steered, the lower portion shows the slope of the 65° isotherm across the Gulf Stream.

---

[2] This survey, as well as the collection of much of the other data presented here, was carried out under Contract No. N6onr–277 with the Office of Naval Research.

MEANDERINGS OF THE GULF STREAM

The May 1946 cruise of the Atlantis agrees with Church's (1937) findings in that the Gulf Stream follows a meandering course after leaving Cape Hatteras. Although only two other short cruises of this type have been made, all of them in the spring of the year, it is apparent at any particular meridian that the Stream may move north or south at the rate of a mile or two a day. At one point, the position of the Stream may be shifted north by 20 miles in about two weeks, while in the same period another part of the Stream may have moved a corresponding distance toward the south. This characteristic of the current makes it impossible to determine by a single section whether or not the Stream as a whole is north or south of its mean position. This point is emphasized, because the north-south migration of the Gulf Stream has been thought to be closely associated with the seasonal variations of the circulation of the North Atlantic.

VELOCITY OF THE GULF STREAM

The data from the May 1946 cruise, and from other more recent cruises of the Atlantis, show that the velocity in the swiftest part of the Gulf Stream is generally between four and five knots. Such high velocities are usually confined to a relatively narrow band 10 to 15 miles wide situated near the inshore or left-hand edge of the current. A number of these calculations of current velocity and direction are plotted in Fig. 4. These calculations and others obtained recently show not only that the highest velocities in the Gulf Stream are greater than most previous estimates, but that currents associated with the eddies both north and south of the Stream are surprisingly strong. The relatively fast (3 knot) currents of these eddies would not often be noticed in general navigation, since for each current moving in one direction there is an opposite current of equal, or nearly equal, velocity near by. It is for this same reason that when crossing the Gulf Stream the ships' drift records show the velocity as less than it actually is. The countercurrents near the Stream are usually indistinguishable from the main flow in the navigational record. Because of the meanderings, eddies and countercurrents which are characteristic of the

Gulf Stream area, it would be difficult for a ship to maintain a course so that it would remain in the swiftest, eastward flowing, portion of the current. Drift records, based on infrequent celestial fixes from a ship following the mean course of the Gulf Stream, would almost certainly not show the maximum velocity of the current. Recently the ATLANTIS spent 36 hours investigating a 3- to 4-knot countercurrent near

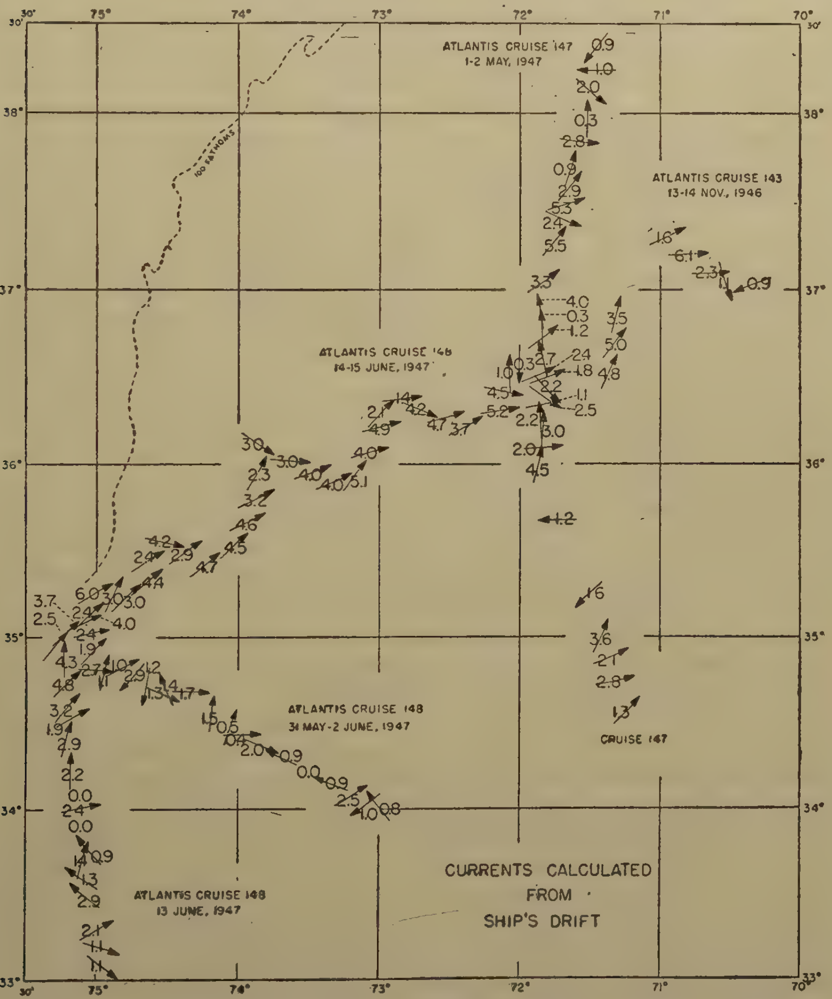
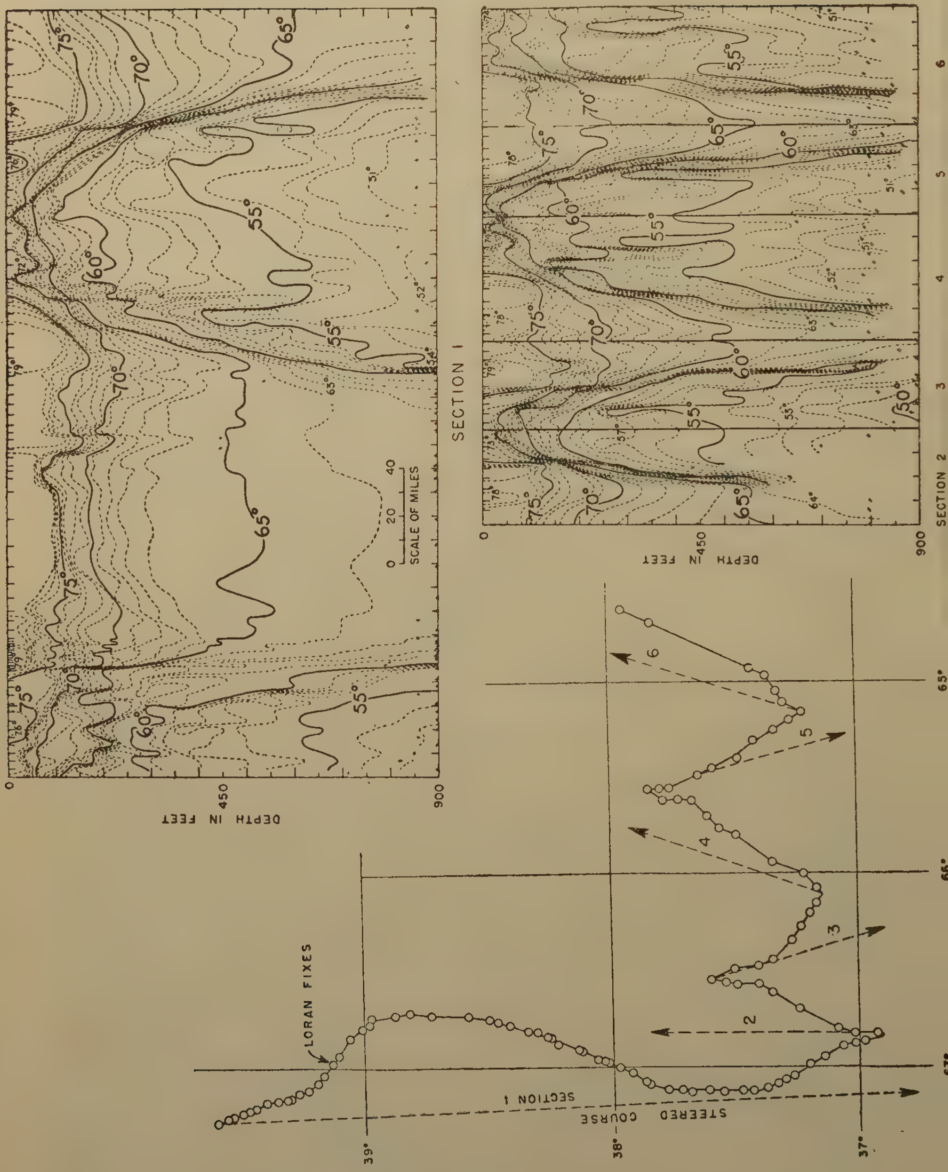

75°
74°
73°
72°
71°
70°
38°
37°
36°
35°
34°
33°
30'
100 FATHOMS
ATLANTIS CRUISE 147
1-2 MAY, 1947
ATLANTIS CRUISE 143
13-14 NOV., 1946
ATLANTIS CRUISE 148
14-15 JUNE, 1947
ATLANTIS CRUISE 148
31 MAY-2 JUNE, 1947
CRUISE 147
ATLANTIS CRUISE 148
13 JUNE, 1947
CURRENTS CALCULATED
FROM
SHIP'S DRIFT

Figure 4. Currents calculated from the departure between the observed and dead reckoning positions during recent Gulf Stream surveys.


SECTION 1
SECTION 2
DEPTH IN FEET
0
450
900
SCALE OF MILES
0
20
40
75°
70°
65°
60°
55°
LORAN FIXES
STEERED COURSE
SECTION 1
2
3
4
5
6
39°
38°
37°
65°

the usual axis of the Gulf Stream, and during this time several freighters bound for Europe were sighted. Since these ships must have been bucking this countercurrent for at least six hours, their calculated drift toward the east for a 12-hour period could not have come to more than one or two knots, even if they had spent the remainder of the time in the swiftest part of the easterly-flowing, true Gulf Stream.

EDDIES ASSOCIATED WITH THE GULF STREAM

The eddies that have been found both north and south of the Gulf Stream fall into two general classes. The *first* are frictionally-driven or shear-zone eddies, often roughly 17 miles in diameter. Spilhaus (1940) found smaller eddies of this type, which he called parasite eddies, having dimensions of the order of magnitude of four miles. The *second* are radically different from these in that they are considerably larger, rotate in the opposite direction, and are frequently found many miles away from the Gulf Stream. These eddies appear to be actual segments of the main stream that have broken away from the parent body when the Gulf Stream, in its meandering, formed an excessively large loop in its path. This would explain the existence of large cyclonic eddies south of the Stream and anticyclonic eddies north of it. The last two sections of the May 1946 cruise of the Atlantis (Fig. 3) showed that the Gulf Stream had then formed a large loop toward the south that could easily have been cut off later by the Stream re-forming across the northern portion. A more fully developed example was observed in June 1947 in the same area. Apparently, we are dealing with a phenomenon that is quite parallel to that observed in the regions of strong westerly winds in the upper atmosphere.

Part of the data from the June 1947 cruise is summarized in Fig. 5. The diagram illustrates the method used in all recent Gulf Stream investigations, that is, the steered course, the frequent loran fixes and the temperature-depth profiles. The relationship between the drift of the vessel (the surface current) and the temperature-depth profile is evident. Sections 2 to 6 along the southern portion of the eddy are typical of profiles crossing the swiftest part of the Gulf Stream itself. As this cruise developed, it became evident that the cyclonic eddy was over 200 miles in length, east-west, and about 60 miles wide. It was partially circumnavigated three times during a period of about eight days.

In 1945, in almost the same area, the schooner Valor had drifted in

---

Figure 5. The navigational record and the corresponding bathythermograph profiles for a crossing of the Gulf Stream south of Georges Banks and also for a large cyclonic eddy (Section 1). Sections 2 to 6 extend along part of the southern edge of this eddy, the center of which was characterized by relatively cold water at all depths.

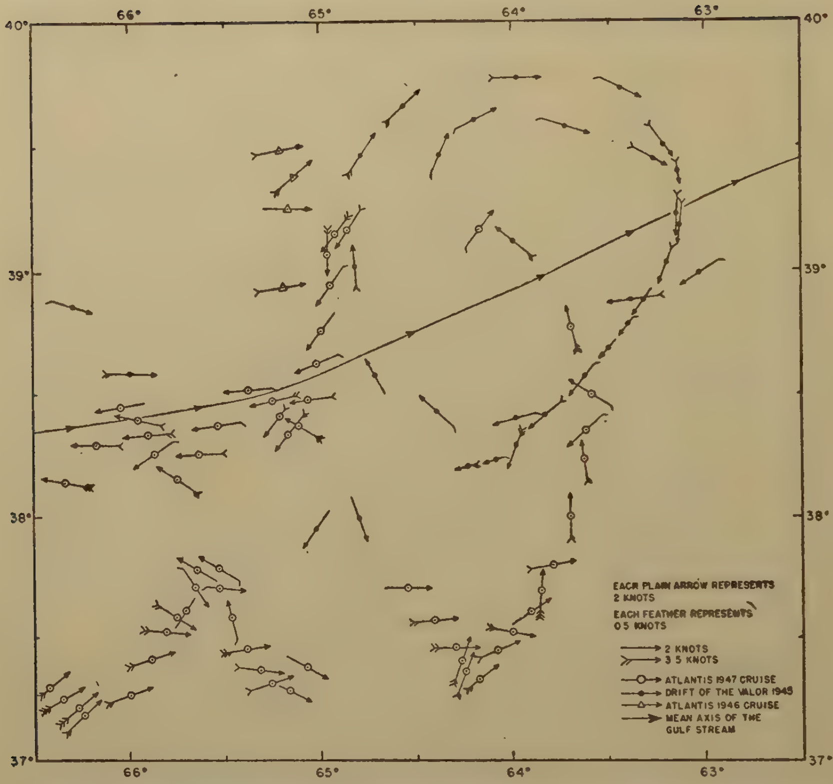

66°
65°
64°
63°
40°
39°
38°
37°
EACH PLAIN ARROW REPRESENTS 2 KNOTS
EACH FEATHER REPRESENTS 0.5 KNOTS
2 KNOTS
3.5 KNOTS
ATLANTIS 1947 CRUISE
DRIFT OF THE VALOR 1945
ATLANTIS 1946 CRUISE
MEAN AXIS OF THE GULF STREAM

Figure 6. Comparison between the currents observed on three recent cruises and the supposed mean track of the Gulf Stream.

a powerful anticyclonic eddy for about 10 days. We are indebted to Lieut. D. T. Parson, USNR, for 48 loran fixes obtained during this period. These, plus his careful records of the drift of the vessel through the water while hove to, have enabled us to make reliable calculations of the velocities encountered. Fig. 6 combines the Valor and Atlantis data. To avoid confusion, only calculated velocities of two knots or more are shown. These observations cover a considerable area through which the Gulf Stream presumably flows normally, yet only a few of the arrows point toward the east. It just happens that on all three occasions the Gulf Stream had been violently displaced; at the time of the 1945 observations it was far to the south; in 1946 the strongest current was about 50 miles north of the mean track. Admittedly, these are probably exceptional observations. The average current charts from which the normal axis of the Gulf Stream has been deduced combine many thousands of less accurate observations. Yet Fig. 6 clearly illustrates the great weakness of such charts. The average current, and the current at any one time, may be very different indeed. Strong currents flowing in deep water are subject to considerable dislocations, and moreover, these strong currents, at least in the western part of the oceans, are considerably narrower than has generally been supposed. If the Gulf Stream is typical of such currents, and it probably is, only by rapid survey techniques can we hope to gain a reasonably reliable picture of its typical behavior. Averaging data merely blurs the details which are the key to its more complete understanding.

BIBLIOGRAPHY

CHURCH, P. E.
1937. Temperatures of the western North Atlantic from thermograph records. Publ. sci. Ass. Océanogr. phys., No. 4: 1–32.
FUGLISTER, F. C.
1947. Average monthly sea surface temperatures of the western North Atlantic Ocean. Pap. phys. Oceanogr. Meteorol., *10* (2): 1–25.
ISELIN, C. O'D.
1940. Preliminary report on long-period variations in the transport of the Gulf Stream. Pap. phys. Oceanogr. Meteorol., *8* (1): 1–40.
NEUMANN, G. AND A. SCHUMACHER
1944. Strömungen und Dichter der Meeresoberfläche vor der Ostküste Nordamerikas. Ann. Hydrogr., Berl., October 1944: 277–279.
ROSSBY, C.-G.
1936. Dynamics of steady ocean currents in the light of experimental fluid mechanics. Pap. phys. Oceanogr. Meteorol., *5* (1): 1–43.
SPILHAUS, A. F.
1940. A detailed study of the surface layers of the ocean in the neighborhood of the Gulf Stream with the aid of rapid measuring hydrographic instruments. J. Mar. Res., *3* (1): 51–75.
STOMMEL, HENRY
1948. The westward intensification of wind-driven ocean currents. Trans. Amer. geophys. Un., *29* (2): 202–206.

SOME EMPIRICAL RELATIONS BETWEEN SEASONAL[1] VALUES OF $(E-P)$ AND SURFACE SALINITIES OVER THE NORTH ATLANTIC[1]

By

WOODROW C. JACOBS
U. S. Weather Bureau,
Washington, D. C.

In the absence of horizontal flow, the surface salinity of any part of the ocean is mainly determined by three processes: Decrease of salinity by precipitation P, increase of salinity by evaporation E, and change of salinity by vertical mixing. Where $(E-P)$ is a positive quantity, the effect on the vertical distribution of salinity is as if salt were being diffused into the volume through the sea surface and, when $(E-P)$ is a negative quantity, the effect is as if salt were being diffused outward through the sea surface. In general, it is to be expected that regions of large $(E-P)$ will be regions of high surface salinities and regions of small or negative $(E-P)$ will be regions of low salinities. Therefore, where the horizontal transport term is small, a quantitative relationship between $(E-P)$ and surface salinity might be expected.

Wüst (1936) has found that the average annual surface salinities can be related to the average annual latitudinal values of $(E-P)$ quite accurately by a simple empirical expression of the form:

$$S_o = K(E-P) + S_k, \qquad (1)$$

where S_o is surface salinity, K is a constant, S_k is a constant salinity and $(E-P)$ is given in cm/year. He finds it necessary, however, to assign different values for K and S_k for different oceans. His formulae do not apply near the coasts where the horizontal transport term tends to be large and where the effects of runoff from land surfaces may be important. They also do not apply near the Equator.

On the basis of previously derived seasonal evaporation and precipitation data for the oceans (Jacobs, 1948), it has been possible to obtain seasonal $(E-P)$ values for the North Atlantic and North Pacific. It is considered of interest to determine whether or not a simple empirical relation can be found between the seasonal salinities and the seasonal quantities of $(E-P)$, similar to the relations found by Wüst, when we consider mean annual values. Unfortunately, seasonal salinity data are not available for the Pacific, while $(E-P)$ values are not available for the South Atlantic because computations of evaporation have not been possible for the latter region. Since the region must be one within which the transport of surface waters is at a minimum, the choice of areas left for analysis is limited.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 400.

An examination of current charts for the oceans (U. S. Hydrographic Office, 1945) indicates that the only completely suitable area is the one in the North Atlantic bounded by the 20th and 30th parallels and the 60° W and 20° W meridians. This section excludes the coastal areas with their pronounced N – S surface currents and is yet of sufficiently large lateral extent to minimize the effects of E – W flow.[2]

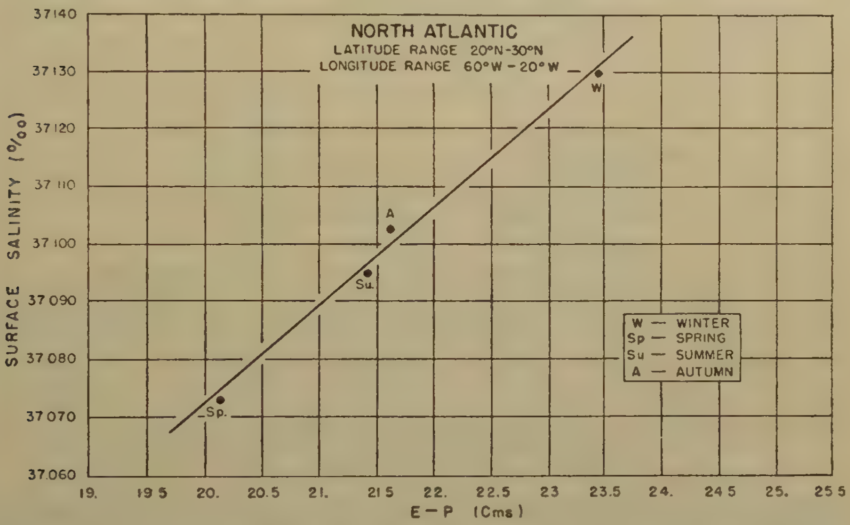


Figure 1. Surface salinity as a function of seasonal ($E - P$) in an area of the North Atlantic where the transport of surface waters is at a minimum (Lat. 20°–30°N; Long. 60°–20° W).

The results of plotting the seasonal salinity values against the seasonal ($E - P$) values for this area are given in Fig 1. An almost perfect straight line relationship is shown such that the surface salinities can be given very accurately by the formula:

$$S_o = 0.0171\,(E - P) + 36.730 \pm 0.003\ (°/_{oo}). \qquad (2)$$

The range of salinities computed through its use differs from the observed range of salinities by less than 10%.

---

[2] The seasonal salinity values have been scaled from charts given by Böhnecke (1936).

Having obtained good results in the one area of the North Atlantic where the transport of surface waters is at a minimum such that the effect of advection on the surface salinity would be small, it was decided to test at least one other area even though a considerable advective effect might be expected. The area chosen for this examination is the one immediately to the north of the previous area, *i. e.*, the area bounded by the 30th and 40th parallels and by the 20th and 60th meridians.

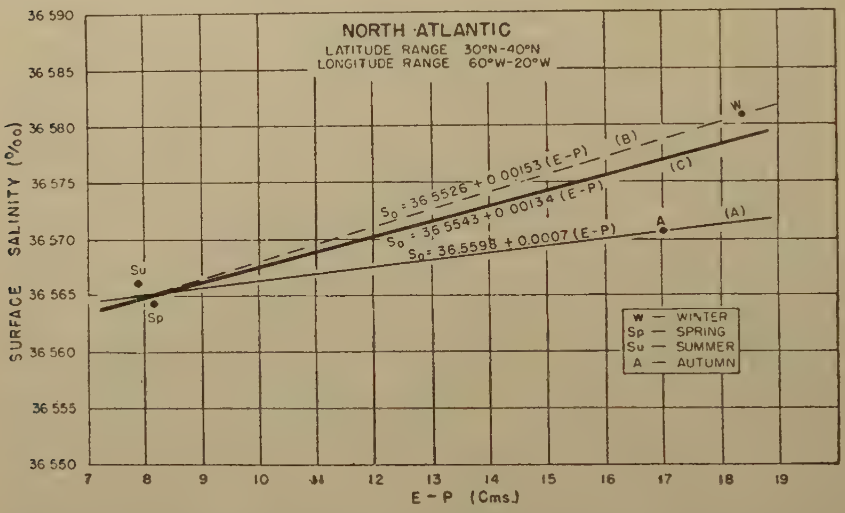


Figure 2. Surface salinity as a function of seasonal $(E-P)$ in an area of the North Atlantic where the horizontal transport of surface waters cannot be neglected (Lat. 30°–40° N; Long. 60°–20° W).

The results of plotting the seasonal values of $(E - P)$ against the seasonal salinities for this area are given in Fig. 2. Unfortunately, the seasonal $(E - P)$ values and seasonal salinities for spring and summer are so nearly identical that, for practical purposes, they both constitute a single point. The winter and autumn points, on the other hand, are so widely dispersed that no reasonable curve can be drawn that would serve to connect all points. However, as first steps in analysis, it is possible to assume that the summer, spring and autumn values are correct and that high salinity water is being advected into the area during winter (curve *A* in Fig. 2) or that the summer, spring and winter values are correct and that low salinity water is being advected into the area during autumn (curve *B* in Fig. 2).

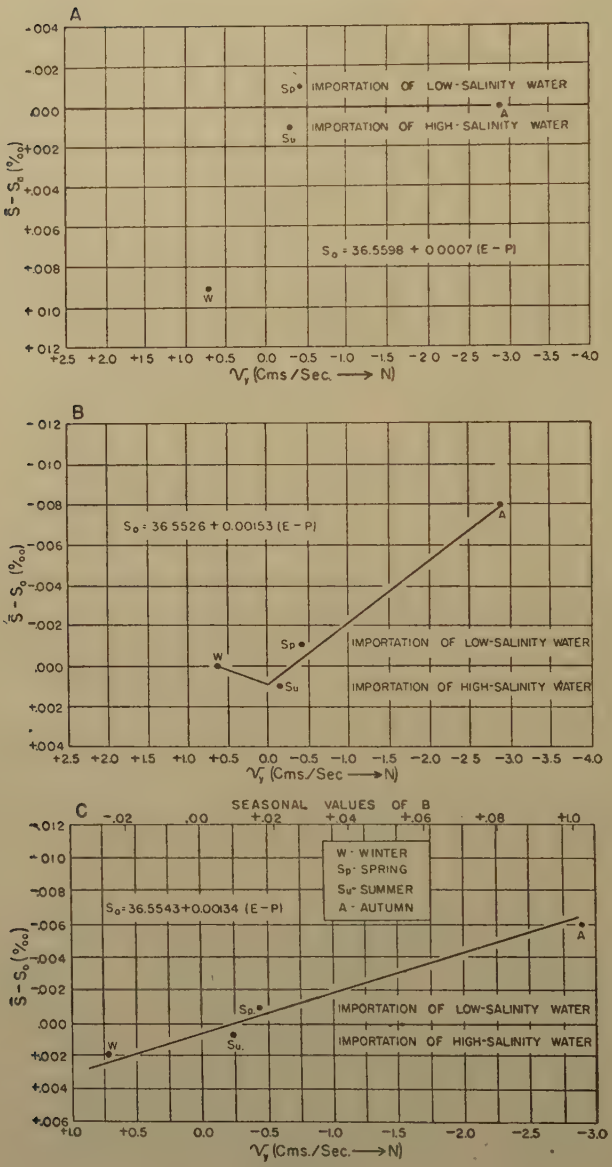
An examination of the salinity distribution over this portion of the North Atlantic, as given by Böhnecke (1936), reveals that the horizontal salinity gradients are slight except along the 40th parallel between the meridians 40° W and 60° W where very low salinity water exists. In order to determine the nature of the surface flow in this area of marked salinity gradients, monthly charts of the surface currents prepared for one-degree squares by the U. S. Hydrographic Office (1945) have been examined and the surface transport across the 40th parallel determined. The midmonths of each of the seasons were accepted as representing the average seasonal flow, and the average northerly component of flow for the area between the 39th and 41st parallels and the 40° W and 60° W meridians was computed for each of these months. The results are as follows:

Season	V_y
Winter	+0.73 cm sec
Spring	−0.43 cm/sec
Summer	−0.26 cm/sec
Autumn	−2.89 cm/sec,

where V_y is the component of flow across the 40th parallel, negative values indicating southerly flow (importation of low salinity water) and positive values indicating northerly flow (transport outward of waters across the 40th parallel) in accordance with usual notational practice.

Given in Fig. 3A are the results of plotting the seasonal values of V_y against the difference between the observed surface salinities, S, and the salinities computed on the assumption that the summer, spring and autumn values are correct (curve A in Fig. 2). An almost complete scattering of points occurs, indicating that the difference between the winter and autumn salinities cannot be accounted for by the advection of high salinity water during the winter season.

In Fig. 3B are plotted the differences between the observed surface salinities and the salinities computed on the assumption that the summer, spring and winter values are correct (curve B in Fig. 2) against the seasonal values of V_y. The three points representing the summer, spring and autumn months fall nicely along a straight line, but the winter value proves to be about 0.003 °/₀₀ too high to be accounted for merely by an absence of advection of low salinity water. However, since the winter season is the only one which shows positive values for V_y (transport toward the north across the 40th parallel), it is reasonable to construct the curve as shown in Fig. 3B by simply connecting the points by two straight lines which intersect at a common point


A
Sp IMPORTATION OF LOW-SALINITY WATER
IMPORTATION OF HIGH-SALINITY WATER
Su
A
W
S₀ = 36.5598 + 0.0007 (E − P)
S̄ − S₀ (‰)
-.004
-.002
.000
+.002
+.004
+.006
+.008
+.010
+.012
+2.5 +2.0 +1.5 +1.0 +0.5 0.0 −0.5 −1.0 −1.5 −2.0 −2.5 −3.0 −3.5 −4.0
V_y (Cms/Sec. ⟶ N)
B
S₀ = 36.5526 + 0.00153 (E − P)
A
Sp
W
Su
IMPORTATION OF LOW-SALINITY WATER
IMPORTATION OF HIGH-SALINITY WATER
S̄ − S₀ (‰)
-.012
-.010
-.008
-.006
-.004
-.002
.000
+.002
+.004
+2.5 +2.0 +1.5 +1.0 +0.5 0.0 −0.5 −1.0 −1.5 −2.0 −2.5 −3.0 −3.5 −4.0
V_y (Cms./Sec ⟶ N)
C
SEASONAL VALUES OF B
−.02 .00 +.02 +.04 +.06 +.08 +1.0
W - WINTER
Sp - SPRING
Su - SUMMER
A - AUTUMN
S₀ = 36.5543 + 0.00134 (E − P)
A
Sp
Su
W
IMPORTATION OF LOW-SALINITY WATER
IMPORTATION OF HIGH-SALINITY WATER
S̄ − S₀ (‰)
-.012
-.010
-.008
-.006
-.004
-.002
.000
+.002
+.004
+.006
+1.0 +0.5 0.0 −0.5 −1.0 −1.5 −2.0 −2.5 −3.0
V_y (Cms./Sec. ⟶ N)

(Figure 3)

along the line where $V_y = 0$ and to make no assumptions regarding the nature of the surface transport of salinity during the winter season.

However, since there is northerly transport of surface waters across the 40th parallel only during the winter season, and considering the fact that, except along the 40th parallel, the salinities of surrounding areas tend to be slightly higher than within the region under consideration, it is more reasonable to assume that water of higher salinity is being advected into the region during winter, the amount advected being proportional also to V_y. On the basis of this assumption the curve *C* in Fig. 2 has been constructed. As shown by Fig. 3C, V_y plotted against the differences between the observed seasonal salinities and the salinities computed on the basis of curve *C* indicate an almost perfect straight-line relationship.

Curve *C* in Fig. 2 can be represented by the following formula:

$$S_o = 0.00134\,(E - P) + 36.5543 \qquad (^\circ/_{oo}), \tag{3}$$

while the curve in Fig. 3C can be represented by the expressions:

$$\bar{S} - S_o = 0.00246 V_y + 0.00071 \qquad (^\circ/_{oo}) \tag{4}$$

or

$$V_y = 406.25(\bar{S} - S_o) - 0.2875 \qquad (\text{cm/sec}). \tag{5}$$

Combining expressions (3) and (4), a simple empirical formula is obtained which allows the surface salinity to be computed with exceptional accuracy. In this case,

$$S_o = 0.00134(E - P) + 0.00246 V_y + 36.555 \pm 0.001 \ (^\circ/_{oo}), \tag{6}$$

where V_y is in cm/sec.

Because of the paucity of subsurface current and salinity data in these regions and because of the complicated nature of the vertical and lateral mixing processes which are involved, it does not appear profitable at this time to attempt to determine, on the basis of theoretical considerations, the significance of the salinity and $(E - P)$ relationships which have been illustrated.

Figure 3. Difference between the computed (S_o) and observed ($\bar{S}$) salinities for the area in the North Atlantic Lat. 30° to 40° N and Long. 60° to 20° W as a function of the transport of surface waters across the 40th parallel between the meridians 60° W and 40° W.

A. The relationship with the northerly component of surface flow, V_y, assuming the relation between $\bar{S}$ and $(E-P)$ as given by curve (*A*) in Fig. 2.

B. The relationship with V_y, assuming the relation between $\bar{S}$ and $(E-P)$ as given by curve (*B*) in Fig. 2.

C. The relationship with V_y, assuming the relation between $\bar{S}$ and $(E-P)$ as given by curve (*C*) in Fig. 2.

The values for the ratio, *b*, between volume of water advected and total volume of water within the mixing layer (assuming V_y constant with depth throughout the mixing layer) are entered along the top margin of graph.

REFERENCES

BÖHNECKE, GÜNTHER

1936. Atlas: Temperature, Salzgehalt and Dichte an der Oberfläche des Atlantischen Ozeans. Wiss. Ergebn. dtsch. Atlant. Exped. *Meteor*, 1925–27, Bd. 5, Atlas: VII + 76.

JACOBS, W. C.

1948. The energy exchange between sea and atmosphere and some of its consequences. 180 pp. ms. to be published.

U. S. Hydrographic Office

1945. Monthly current charts—North Atlantic. H. O. Misc. No. 10,688-1-.

WÜST, GEORG

1936. Oberflächensalzgehalt, Verdunstung und Niederschlag auf dem Weltmeere; In, Länderkundliche Forschung, Festschrift Norbert Krebs, pp. 347–359. J. Engelhorns Nachf., Stuttgart. 368 pp.

FOG DEVELOPMENT AT SAN DIEGO, CALIFORNIA

By

DALE F. LEIPPER

Scripps Institution of Oceanography[1]
University of California
La Jolla, California

INTRODUCTION

If true fog alone be considered—that is, suspended water particles which reduce horizontal visibility at the earth's surface to 1,000 meters (or ½ mile) or less—then fog at San Diego occurs nearly four times as frequently during the winter period, from October through April, as during the remaining months of the year (11, p. 34). Yet, of the many studies of California fog which have been made, none appears to be concerned directly with coastal fog in the winter period.

The fog situations during several winter periods were examined at Scripps Institution. Particular attention was paid to air mass characteristics at various altitudes relative to sea surface temperature. It became apparent that nearly all of the fog situations developed in a systematic manner over a period of days, and that the developments resulted from the effect of local influences upon a given set of initial conditions. The purpose of the following discussion is to describe this development, to present a model, and to give statistical evidence showing the pertinence of the concepts involved.

A MODEL DEVELOPMENT

A model has been devised to illustrate the manner in which fog situations develop in the San Diego area. The model is described in four different stages beginning with the initial conditions which are required if the development is to take place.

The dominant feature of the initial conditions is the presence over the sea of air which has a temperature higher than the sea surface temperature and which is relatively dry aloft. The high temperature of the air guarantees that an inversion will be formed at the surface to restrict the vertical movement of moisture and thus cause a thin lower layer of air to approach saturation. The dry air aloft permits rapid cooling of the surface layer and intensification of fog by radiation.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 381. The work represents results of research carried out for the Office of Naval Research, Navy Department, under contract with the University of California.

The initial conditions for fog formation may be brought about when the Pacific anticyclone pushes inland over northern California and causes a general easterly flow, so that the air arriving at San Diego has moved down the slope of 4,000-foot mountains. Air following such a path is necessarily warm and dry at San Diego. Out of 42 winter fogs occurring in 1943 to 1945, 39 were preceded or accompanied by easterly winds aloft. The typical synoptic weather situation is shown in Fig. 1.

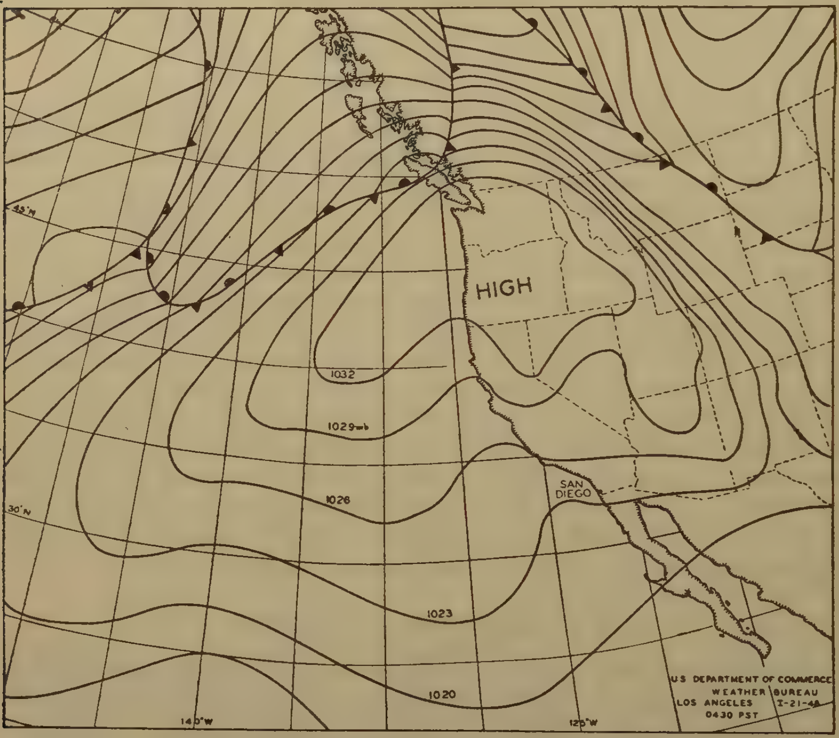


Figure 1. Weather map showing the type of synoptic situation which usually precedes winter fog at San Diego.

In the second stage of the fog development, the easterly air flow weakens as the high-pressure cell is divided into two parts, one on each side of the coastal mountains. Flow of warm dry subsiding air toward the sea decreases, and the locations of the high-pressure cells permit the air offshore from San Diego to remain in a relatively stagnant condition. The moisture content of the lower layers increases due to evaporation from the sea, and the heat content of these layers decreases due to conduction of heat downward (9). A surface inversion is formed, and the lowest air layers become nearly saturated at a temperature close to that of the underlying sea surface.

In the third stage, the normal northwesterly airflow and seabreeze regime returns and fog forms. The details of formation cannot be described as yet, but observations indicate that formation occurs over regions of low sea surface temperature. Fig. 2 shows a tongue of relatively cold water which is consistently found in the same area. Fog formation is often observed here (1). Also, fog is often found along the coastline where observations during eight winter cruises by the buoy boat of the Scripps Institution have shown a band of water which is at times several degrees colder than the sea surface ten to

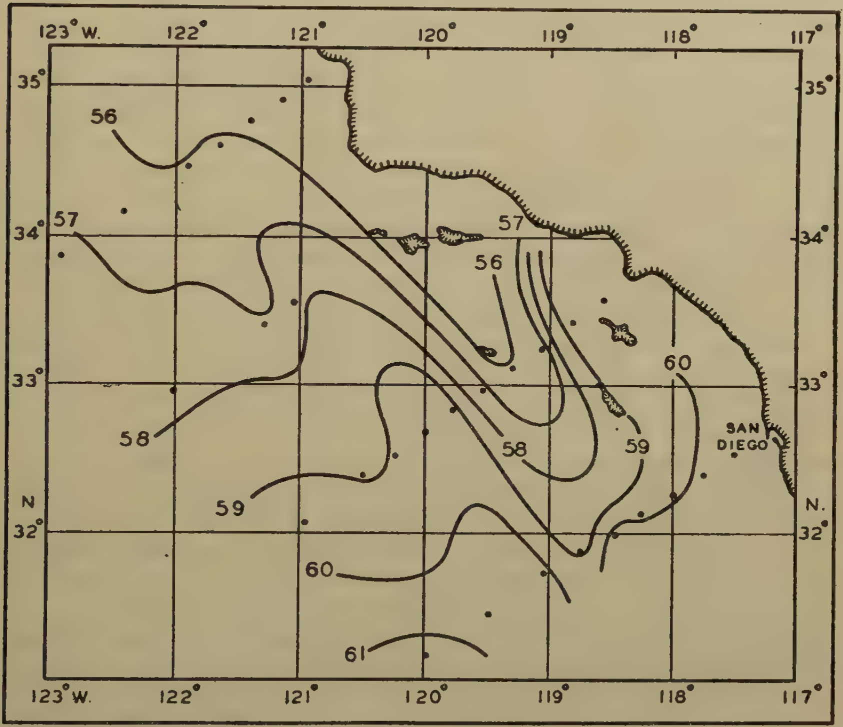


Figure 2. Sea surface isotherms in degrees Fahrenheit based upon a cruise of the E. W. Scripps, February 15 to 25, 1938. Observations were made where dots appear.

fifteen miles offshore. In the situation shown in Fig. 2, the average sea surface temperature at Scripps pier near San Diego was 14.3C (57.8 F) and that offshore was 15.6C (60.0 F).

Changes in humidity and heat content which occur during flow of nearly saturated air across sea surface isotherms may cause the formation of fog. If these changes do not form fog, they appear to bring the thin lower layer of air into such a state that large amounts of heat are lost by radiation from the moist air, and fog is formed by this means (4).

Once fog is formed, radiation from the cloud cools the thin, nearly stagnant fog layer to a temperature several degrees lower than the sea surface temperature. Under these conditions there is continual upward flux of heat at the sea surface, but the air remains colder than the sea. The lapse rate of temperature is nearly dry adiabatic (7). Since the air is colder than the sea, evaporation continues even though the water vapor is saturated (9). Thus, radiational cooling causes first a decrease in temperature and then an increase in evaporation. Both of these effects markedly intensify fog and serve to maintain it.

As the sea breeze strengthens during the daytime, the fog is carried shoreward and is likely to arrive at the coast in the late morning or early afternoon. The day on which it will first arrive is determined by the time at which the pressure gradient, which caused the initial easterly flow, weakens sufficiently to permit a sea breeze to develop.

When the fog layer first arrives at the coast it is usually quite thin and is soon dissipated by the daytime heat of the land; thus the fog does not extend inland. Since the temperature at the surface increases rapidly as the air moves inland, the moisture is carried upward and the dew point inland from the coast does not increase materially. At night, the sea breeze weakens, and moisture is no longer brought toward land. Thus, there is not sufficient moisture for fog to be formed inland either during the day or during the night.

Stage four of the fog development occurs when the nearly adiabatic layer deepens to approximately 400 feet or more. The increasing westerly winds and sea breezes and the convection resulting from radiational cooling at the cloud top contribute to this deepening. Evaporation from the sea continues and the presence of the strong inversion still restricts the vertical movement of moisture, thus maintaining a high dew point in the marine layer (the nearly adiabatic moist layer). Fog is present over the sea both day and night. During the day the sea breeze carries the fog toward shore. The daytime radiation and conduction of heat from the land mass dissipate the fog as it approaches shore, but the nearly adiabatic layer, being deeper than in the previous stage, is not destroyed as rapidly. It retains its high moisture content in the form of vapor and moves inland. At night, the air over the land cools and fog is formed in the moist layer. As cooling continues, this fog may spread and join with the fog offshore.

As the marine layer below the inversion continues to deepen, the point is reached where the mixing and cooling processes do not result in condensation throughout the entire layer. The upper portion is filled with water particles first and fog does not form on the ground. Thus, the fog sequence is ended and the stratus regime begins. The complete fog sequence usually extends over a period of about five days.

In the actual development of a particular situation, deviations from the above transformation may be observed. For example, on rare occasions the entire sequence may take place in a period as short as one day; a normal sequence may be interrupted or terminated at any stage by the passage of a front or an upper pressure trough; or, after the easterly winds have weakened and fog development has started, the winds may strengthen again, thus temporarily reversing the development. In spite of the deviations of particular cases from the model, nearly every case of fog in the winter period may be identified with the development described above.

AN OBSERVED DEVELOPMENT

In Fig. 3 are shown the soundings for seven successive days in January 1948. Sounding A shows the presence of dry air warmer than the sea to heights as great as 2,000 feet, which indicates the existence of stage one described in the model development. This is a warning of fog to come. The synoptic situation associated with the influx of this air is similar to that of Fig. 1, the time of the map being nine and one-half hours after the time of the sounding. Sounding A is an afternoon sounding which was used because the morning observation was not available on this day. The remainder of the soundings in Fig. 3 are early morning observations made at 24-hour intervals.

Sounding B shows that a surface inversion has formed and that the surface air temperature is lower than the sea surface temperature. This indicates that stage two occurred between soundings and that stage three exists. Fog arrived at the coastline at 1710 PST (Standard Time at 120° W) on the afternoon following the sounding, January 21. The air temperature in the fog as it arrived at the end of the Scripps pier, 16 miles north of San Diego, was 2.4 C lower than the sea surface temperature there. It was also lower than any known sea surface temperature in the near-by offshore area. Fog was observed at Lindbergh Field, San Diego, for 48 minutes beginning at 1820 PST. It retreated when the sea breeze, which had been WNW at 3 to 7 miles per hour, weakened and was replaced by S and E winds of 2 mph. The breakdown of the Pacific anticyclone into two cells occurred prior to 2235 PST, as shown by the synoptic weather map at that time.

Sounding C indicates the presence of a mixed layer below the inversion and is associated with fog at Lindbergh Field in the early morning and for a short period in the early evening. Careful measurements of sea surface and air temperatures were made at approximately half-hour intervals in the fog at distances 0.2 to 5.0 miles from shore on this day. Air temperature was observed at heights of a few inches, five feet, thirty feet, and seventy feet above the sea surface. Between 0950 and 1317 PST, when the observations were made, every air temperature was lower than the sea surface temperature. The average difference was 2.3 C.

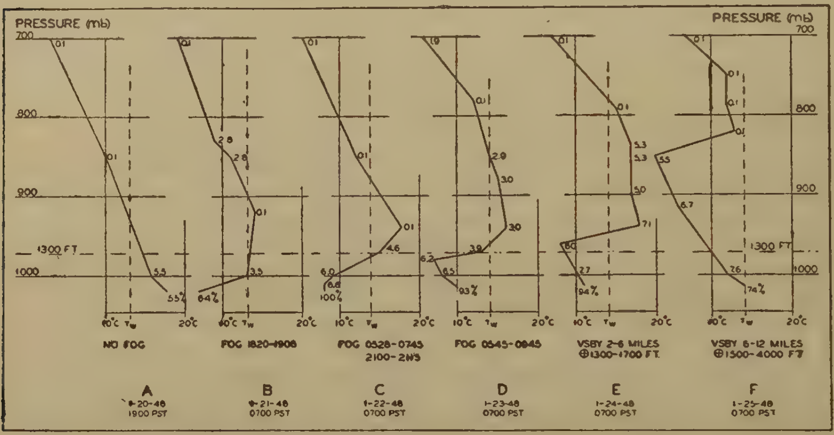


Figure 3. Soundings during the development of a fog situation, taken from radiosonde observations at North Island, San Diego, California. The number at the base of each sounding is the relative humidity. Other numbers along the soundings are specific humidities in grams per kilogram. Vertical lines are isotherms labeled at the base, and the dashed line, labeled T_w, is the sea surface temperature.

In sounding D the mixed layer has deepened but is still only 1,000 feet thick. On this day, which is typical of stage four in the model development, fog at Lindbergh Field formed at 0545 PST and lasted four hours. On the following day, January 24, the mixed layer became too deep for fog to form; visibilities were from two to six miles, and an overcast persisted all day with ceiling height from 1,300 to 1,700 feet. Thus the sequence ended. On the day represented by sounding F the highest temperature above the inversion had dropped below the sea surface temperature and the inversion base was too high to permit fog formation. Visibilities continued to improve and were from six to twelve miles. The sky remained overcast with ceilings from 1,500 to 4,000 feet.

NONDIURNAL INDICES OF FOG

In order to characterize the stage of the fog situation development where fog may be expected at Lindbergh Field, three nondiurnal indices have been defined. Each element used in determining values of the indices is measured at a given time of day. Thus diurnal variations do not affect the index values. So that a day would usually include both the formation and dissipation time of fog, a 'fog day' was defined as the period from 1630 PST on one calendar day through 1629 on the following calendar day. The indices for a given fog day are:

1. Height of the Inversion Base: The height above which the air temperature increases with height at the most rapid rate on the North Island (San Diego) morning raob (radiosonde observation).
2. Temperature Index: The quantity $(T_a - T_w)$ where, if an inversion exists with base below 3,000 feet, T_a is highest air temperature above the inversion base on the morning raob or, if no inversion exists with base below 3,000 feet, T_a is the surface air temperature on the morning raob; T_w is the sea surface temperature at the end of the Scripps pier.[2]
3. Moisture Index: The quantity $(DP_{1630} - T_w)$ where DP_{1630} is the surface dew point at Lindbergh Field at 1630 PST and T_w is the sea surface temperature on the preceding fog day.

Values of the three nondiurnal fog indices have been determined for each day of the period 1943 through 1945, when data were available. It was observed that most days on which fog occurred were days when the indices had values lying within certain ranges. Limits were placed on these ranges after a study of the variations. The ranges, determined to be favorable to fog, were: for height of inversion base, 0 to 1,300 feet; for temperature index, any positive value; for moisture index, any positive value or any negative value between 0 and -5 C.

The index ranges selected as favorable to fog in the winter period have been found to be applicable also to fogs which develop in the summer period. However, in summer the synoptic weather features which bring dry warm air over the sea and cause the inversion base to lower into the range favorable for fog may be different than in winter. Blake (2) has shown large-scale zonal subsidence to be an important process.

---

[2] Regular sea surface temperature observations are not made closer to San Diego than the Scripps pier; however, selected observations have shown that the temperature at the pier is usually representative of the coastal regions closer to San Diego and in San Diego Bay.

In the period 1943–1945 there were 222 days when all three indices were favorable to fog; on 62 of these days (28%) fog (*i. e.*, visibility ½ mile or less) occurred. There were 825 days which did not favor fog because at least one index was unfavorable; on 23 of these days (2.8%) fog occurred. Thus, the percentage of fog occurrences on days defined by the indices as favorable to fog was ten times as great as the percentage of occurrences on unfavorable days. Furthermore, on the 23 days when fog occurred with unfavorable conditions existing at normal times for the determination of index values, linear interpolation (or allowance for observational error) shows that the indices were probably favorable at the time of fog formation. There was one exception, when a period of fog lasted only 40 minutes; this was associated with rapidly falling pressure and was followed by rain and strong frontal activity.

For practical purposes it may be stated that fog will not form at Lindbergh Field on a given day unless the three nondiurnal indices are favorable. These indices thus provide a necessary but not sufficient condition for fog formation. If the conditions are to become sufficient, other considerations must be employed. One which shows considerable promise is an evaluation of the effect of moisture in the atmosphere above the inversion base. In 1943, for example, there were 96 days favoring fog and 32 days when fog occurred. In that year, one when reliable moisture data were available, it was found that 47% of the 64 days favorable to fog but with no fog occurring were days when the moisture content aloft was high, as indicated by the mixing ratio at 10,000 feet being 3.5 grams per kilogram or greater. If 1943 is a typical year, then fog will form on 48% of the days having the three nondiurnal indices favorable and also having a mixing ratio at 10,000 feet less than 3.5. On the majority of the remainder of such days there will be a low overcast or a fog bank at sea.

The indices have been plotted against time over a three-year period. It appears that they vary in a relatively slow and consistent manner which may be forecast, and that they may be of value in day-by-day fog forecasting for Lindbergh Field. Since fog formation, wherever it may occur, is controlled by some of the same factors and processes which operate at San Diego, it may be possible to apply similar reasoning and indices to other locations.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Dr. H. U. Sverdrup for encouragement and support in carrying out the work upon which this discussion is based. Mr. Dean Blake gave helpful advice based upon his forty-five years' experience in weather problems in the San Diego area. The local office of the U. S. Weath. Bureau, the Naval Air Station, and the Navy Electronics Laboratory have furnished data and valuable co-operation.

SUMMARY

A model illustrating the manner in which fog situations develop in the winter period in the San Diego area is presented. The importance of the role played by sea surface temperature is stressed. An observed example of the development is described. Nondiurnal indices are defined, these serving as criteria to aid in determining whether or not fog will form at Lindbergh Field, San Diego, on a given day in either summer or winter.

BIBLIOGRAPHY

1. Anderson, J. B.
    1931. Observations from airplanes of cloud and fog conditions along the southern California Coast. Mon. Weath. Rev., Wash., *59* (7): 264–270.
2. Blake, Dean
    1948. The subsidence inversion and forecasting maximum temperature in the San Diego area. Bull. Amer. met. Soc., *29:* 288-293.
3. Byers, H. R.
    1930. Summer sea fogs of the Central California Coast. Univ. Calif. Publ. Geogr., *3* (5): 291–338.
4. Emmons, Gardner and R. B. Montgomery
    1947. Note on the physics of fog formation. J. Meteorol., *4* (6): 206.
5. Leipper, D. F.
    1947. California stratus forecasting correlations, 1935 and other years. Bull. Amer. met. Soc., *29:* 294-297.
6. Neiburger, Morris, C. G. P. Beer and L. B. Leopold
    1945. The California stratus investigation of 1944. U. S. Dept. Commerce, Weath. Bur. 83 pp.
7. Petterssen, Sverre
    1938. On the causes and the forecasting of the California fog. Bull. Amer. met. Soc., *19* (2): 49–55.
8. 1939. Some aspects of formation and dissipation of fog. Geofys. Publ., *12* (10): 1–22.
9. Sverdrup, H. U.
    1942. Oceanography for meteorologists. Prentice-Hall, Inc., New York. 246 pp.

10. Taylor, G. I.
    1917. The formation of fog and mist. Quart. J. R. met. Soc., *43* (183): 241–268.
11. U. S. Weather Bureau Climatologists
    1944. Fog and low ceiling frequencies along the U. S. Pacific Coast. Weath. Div., Hq. Army Air Forces, Rep. No. 30 (rev.). 52 pp.
12. Willett, H. C.
    1928. Fog and haze, their causes, distribution, and forecasting. Mon. Weath. Rev., Wash., *56* (11): 435–468.

THE EXCHANGE OF OXYGEN ACROSS THE SEA SURFACE[1]

By

ALFRED C. REDFIELD

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts
and
Department of Biology, Harvard University
Cambridge, Massachusetts

In temperate latitudes the oxygen content of sea water near the surface varies with the season. This results because the solubility of oxygen in the water changes with its temperature and because of the unequal production and utilization of oxygen by organisms at different seasons. Both factors cause an exchange of oxygen across the sea's surface. The present paper is an attempt to determine the magnitude of this exchange as it occurs in the Gulf of Maine and to evaluate the factors responsible.

The solution of the problem depends upon the assumption that the column of water underlying a square meter of sea surface and extending to a depth beyond which seasonal changes in chemical composition are insignificant may be treated as a closed system except for exchanges with the atmosphere. If Q_o represents the quantity of dissolved oxygen observed to be present in the column and ΔQ_o its change between two periods of observation, ΔQ_E the quantity entering through the sea surface, and ΔQ_x the net increase in oxygen resulting from biological activity during the time, then the equation representing the net exchange across the sea surface is

$$\Delta Q_E = \Delta Q_o - \Delta Q_x. \tag{1}$$

Negative values of ΔQ_o and ΔQ_x represent decreases in the dissolved oxygen in the column and negative values of ΔQ_E represent oxygen escaping to the atmosphere.

EVALUATION OF ΔQ_o AND ΔQ_x

The oxygen content of the column, Q_o, may be obtained by integrating the area enclosed by a curve representing the station data for oxygen concentration and depth. From the difference in Q_o obtained at successive surveys ΔQ_o is estimated.

---

[1] Contribution No. 442 from the Woods Hole Oceanographic Institution.

The net change in oxygen content due to biological activity, ΔQ_x, may be deduced from observed changes in inorganic phosphate content of the column, PO_4. In the process of photosynthesis a quantity of oxygen is set free proportional to the phosphate withdrawn from the water. Conversely, the oxidation of organic matter in metabolism sets free a quantity of phosphate corresponding to the oxygen consumed (Redfield, 1934; Fleming, 1941; Edmondson and Edmondson, 1947). In order to convert changes in phosphate into the quantities of oxygen set free or absorbed in the accompanying biological processes, use is made of an equivalence based on the observation (made in the waters of the Sargasso Sea) that the disappearance of 1 cc O_2 accompanies the appearance of 0.373×10^{-3} mg-at of phosphate phosphorus (Redfield, 1942)[2]. Correspondingly, the disappearance of 1 mg PO_4 is accompanied by the appearance of 28.3 cc O_2, or

$$\Delta Q_x = 28.3 \quad PO_4 .$$

Some uncertainty exists regarding the exactness of this relation in its present application, since it depends upon the assumption that in decomposition the process has gone to completion. Actually significant quantities of phosphorus are present in sea water at certain seasons in the form of dissolved organic compounds (Redfield, Smith and Ketchum, 1937). These do not contribute to the measurement of phosphate concentration; yet it cannot be doubted that their liberation was attended by some oxidation of associated organic matter. Increases in inorganic phosphate may consequently lag behind the attendant consumption of oxygen at the times and places where decomposition is occurring. Moreover, the value of the relation depends upon the average elementary composition of the organisms responsible, which may differ somewhat with time and place. Finally, the methods for determination of inorganic phosphate are not wholly satisfactory. Estimations of ΔQ_x are consequently subject to much greater errors than those of ΔQ_o.

DATA

The data examined were collected in the course of a survey of the Gulf of Maine made by the ATLANTIS in 1933–34. Eleven cruises were completed in the course of 15 months.

The area chosen for examination was the western basin of the Gulf, a region of 200 to 250 m in depth lying between Cash's Ledge and the Massachusetts Coast. Within an area 35 miles square, lying between the latitudes 42° 15′ and 42° 50′ N and the longitudes 69° 10′ and 69° 45′ W, there were available from two to four sets of station data for each cruise[3].

---

[2] This equivalence is based on phosphate data uncorrected for salt error. See Cooper (1938).

The data were collected as part of a general hydrographic survey, and consequently no special provision was made to secure measurements designed to meet the special requirements of the present study.

In evaluating the quantity of oxygen and phosphate in the column under a square meter of surface at each period of observation, the data from the several stations occupied in the area were averaged. Oxygen data were available from three stations for each cruise except for January and August, when two stations were occupied, and for December, when four stations were occupied in the area. Phosphate data were available from three stations in September 1933, from two in July and December 1933 and in March and May 1934, and from one in January, April, June and September 1934. No data were available for October 1933 and August 1934.

It was assumed that seasonal changes in the quantity of oxygen in the water column are limited to a depth of 200 m and of phosphate to a depth of 100 m. Oxygen and phosphate contents were consequently integrated to these depths in estimating ΔQ_o and ΔQ_x.

The content of dissolved oxygen and of inorganic phosphate were both maximal in March. Consequently, it proved convenient to express ΔQ_o and ΔQ_x as the changes which have occurred or which will occur relative to the conditions obtaining at that time.

ASSUMPTIONS

The basic assumptions underlying the application of equation (1) are that exchanges of oxygen across the sides and lower surface of the water column have a negligible effect on its oxygen content and that equation (2) is valid. The latter assumption is sufficiently discussed above and in the references.

The assumption that exchanges resulting from the movement of oxygen or phosphate through the side walls of the column are negligible can be justified only by the consideration that the area of observation is centered in a much larger region in which the horizontal distribution of components is uniform at any depth. The components which leave the column from one side by lateral mixing or flow may be supposed to be replaced almost exactly by those which enter from the opposite side. The seasonal changes in the concentration are so large compared to the gradients which are present in a horizontal direction at any time

---

[3] For data see Bulletin Hydrographique 1933 and 1934 M/Y Atlantis Stations Nos. 1699, 1703, 1721, 1747, 1748, 1749, 1856, 1857, 1858, 1864, 1865, 1875, 1876, 1909, 1910, 2022, 2023, 2024, 2074, 2075, 2076, 2168, 2169, 2170, 2220, 2221, 2236, 2258, 2259, 2273, 2274, 2275. Hydrographic observations were made by standard procedures. Oxygen was determined by Winkler's method and phosphate by the Denigès-Atkins procedure. The values for phosphate are not corrected for salt error.

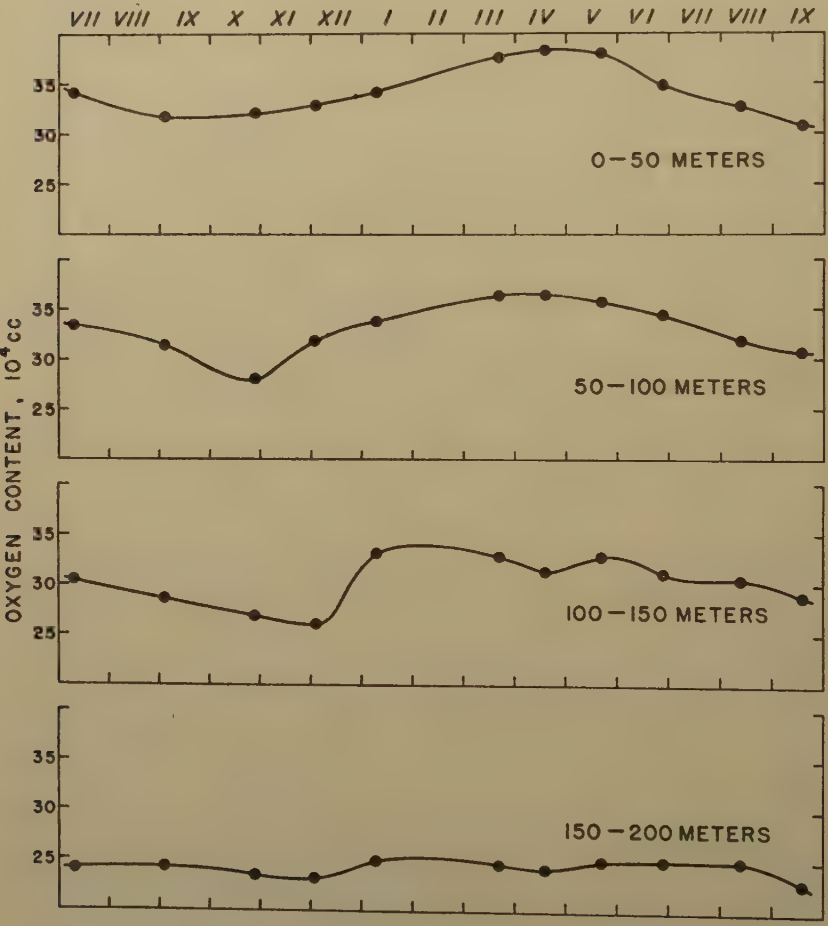


Figure 1. Seasonal variation of oxygen content of successive zones of 50 m depth beneath one m^2 of sea surface in Gulf of Maine.

that it seems unlikely that lateral transport has an important effect on the changes in the composition of the column.

The assumption that exchanges through the lower surface of the column are negligible requires

(1) that its depth be greater than that at which seasonal effects on the components can be observed and

(2) that the net movement of water from below by vertical mixing is insignificant.

The western basin of the Gulf of Maine is not sufficiently deep to permit the choice of a depth greater than 200 m for consideration. An examination of the samples collected periodically at greater depths shows only small fluctuation in oxygen content, and these do not appear to be related to the seasonal fluctuations observed at lesser depths. As a further check, the oxygen content of each 50-m length of the column has been estimated separately and the results plotted in Fig. 1. In the zone between 150 and 200 m the annual change in oxygen content is very small and amounts to only 10% of that of the entire 200-m column. Evidently failure to take account of seasonal changes which may occur at greater depths does not introduce significant errors.

A similar treatment of the phosphate data is shown in Fig. 2. The curve for the zone extending from the surface to 50 m shows a large seasonal fluctuation, which accounts for about 60% of the total seasonal change. The zone between 50 and 100 m also shows an increase in phosphate during the winter period. In the zones between 100 and 200 m there are fluctuations less clearly related to the general seasonal cycle and produced chiefly by low values obtained in January and April. Since there is no reason to think that phosphate concentrations would change in the deeper waters in this way, it is suspected that these fluctuations are due to systematic errors in the phosphate determinations during the cruises of January and April, such as might be occasioned by faulty standards. In order to avoid the multiplication of such errors, it has been assumed in the following consideration that the seasonal changes in phosphate below 100 m are negligible and that the best estimate of the changes in oxygen content due to biological activity are to be obtained from the phosphate content of the upper 100 m. Neglect of the changes in phosphate in the zone between 100 and 200 m reduces the estimated maximal seasonal variation of the column by only 20% and does not alter materially the general conclusions reached by this study.

Vertical mixing probably causes some exchange of oxygen and phosphate across the lower boundary. The deeper water of the Gulf has its origin through an inflow of slope water at the Eastern Channel, and this water becomes fresher and colder as the western basin is approached, effects which can only be attributed to admixture with the overlying layers (Bigelow, 1927). It seems likely that there is a net upward movement of water as the result of this vertical exchange, in which an excess of deep water rises to escape from the Gulf in the surface currents across Georges Bank. The magnitude of this exchange cannot be evaluated and the phenomena require further study.

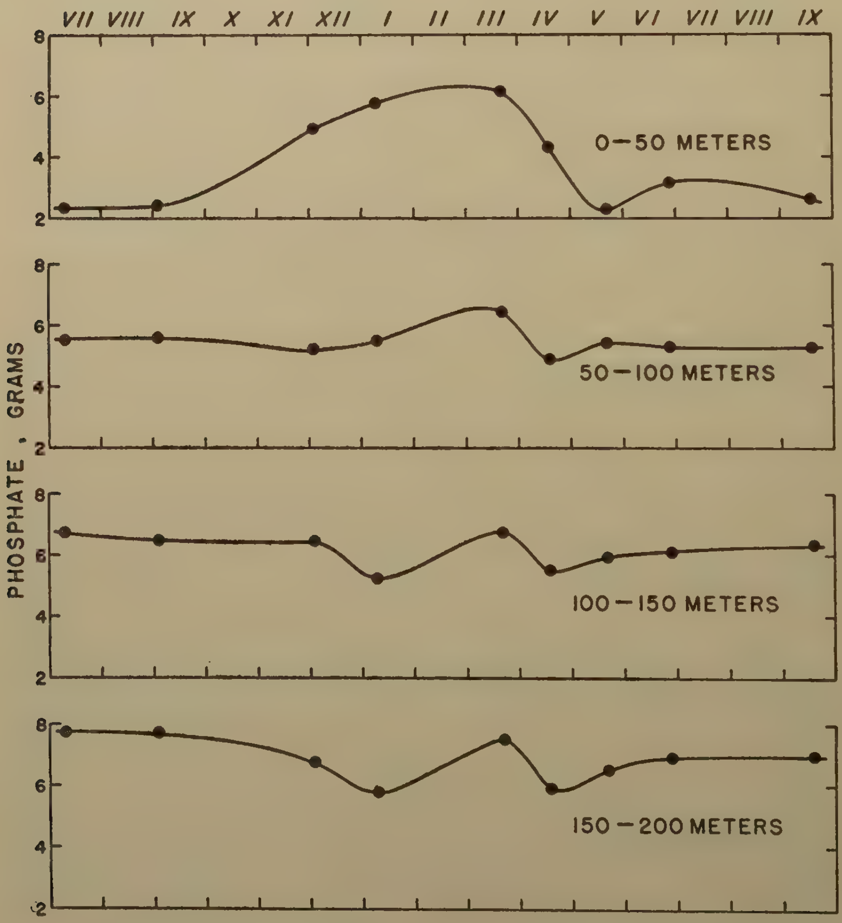


Figure 2. Seasonal variation of inorganic phosphate content of successive zones of 50 m depth beneath one m² of sea surface in Gulf of Maine.

It is likely, however, that such exchanges are small compared to the seasonal effects under consideration.

RESULTS

Table I records the results of the computations. The values obtained for ΔQ_o, ΔQ_x, and ΔQ_E are shown plotted against time in Fig. 3.

TABLE I. Oxygen and Phosphate Content of Water beneath One Square Meter of the Surface of the Gulf of Maine and the Seasonal Changes Estimated to Occur Relative to the Conditions Obtaining in March

Period of Observation	*1* Q_o *10⁴ cc*	*2* ΔQ_o *10⁴ cc*	*3* PO_4 *g*	*4* ΔPO_4 *g*	*5* ΔQ_x *10⁴ cc*	*6* ΔQ_E *10⁴ cc*
July 6–10, '33	122.2	−9.3	7.79	−4.86	13.7	−23.0
Sept. 4, '33	115.9	−15.6	8.17	−4.48	12.6	−28.2
Oct. 28–29, '33	110.4	−21.1	——	——	(10.0)	(−31.1)
Dec. 2–4, '33	113.8	−17.7	10.17	−2.48	7.0	−24.7
Jan. 9, '34	126.1	−5.4	11.27	−1.38	3.9	−9.3
March 22, '34	131.5	0	12.65	0	0	0
April 18, '34	130.1	−1.4	9.18	−3.47	9.8	−11.2
May 22, '34	131.3	−0.2	7.77	−4.88	13.8	−14.0
June 26–July 1	125.1	−6.4	8.44	−4.21	11.9	−18.3
August 11, '34	119.9	−11.6	——	——	(12.5)	(−24.1)
Sept. 19, '34	113.0	−18.5	7.89	−4.76	13.4	−31.9

Column *1* Q_o, oxygen content of column 1 m² × 200 m deep.
2 ΔQ_o, change in Q_o relative to March.
3 PO_4, inorganic phosphate content of column 1 m² × 100 m deep.
4 ΔPO_4, change in PO_4 relative to March.
5 ΔQ_x, change in oxygen content, corresponding to ΔPO_4, attributed to biological activity. It is assumed that ΔPO_4 and ΔQ_x are negligible below 100 m.
6 ΔQ_E, exchange of oxygen across 1 m² sea surface, relative to March.
Values in parentheses obtained by interpolation from curves in Fig. 3.

The annual change in the oxygen content of the water (maximum numerical value of ΔQ_o) is about 20 × 10⁴ cc per m² of sea surface. The content is maximal in the latter part of March and minimal at the end of October.

The net annual change in oxygen content attributable to biological processes (maximum value of ΔQ_x) is about 13 × 10⁴ cc per m² of sea surface. Oxygen production exceeds consumption during April and May. The processes are nearly balanced from June to September. Consumption exceeds production between October and March.

The annual exchange of oxygen across the sea surface (maximum numerical value of ΔQ_E) is about 30 × 10⁴ cc per m². This quantity leaves the sea water during spring and summer and re-enters during the winter seasons. The turning points occur in the latter parts of October and March. The quantity of oxygen involved in the exchange is equal to that in a layer of air 1.5 m in depth.

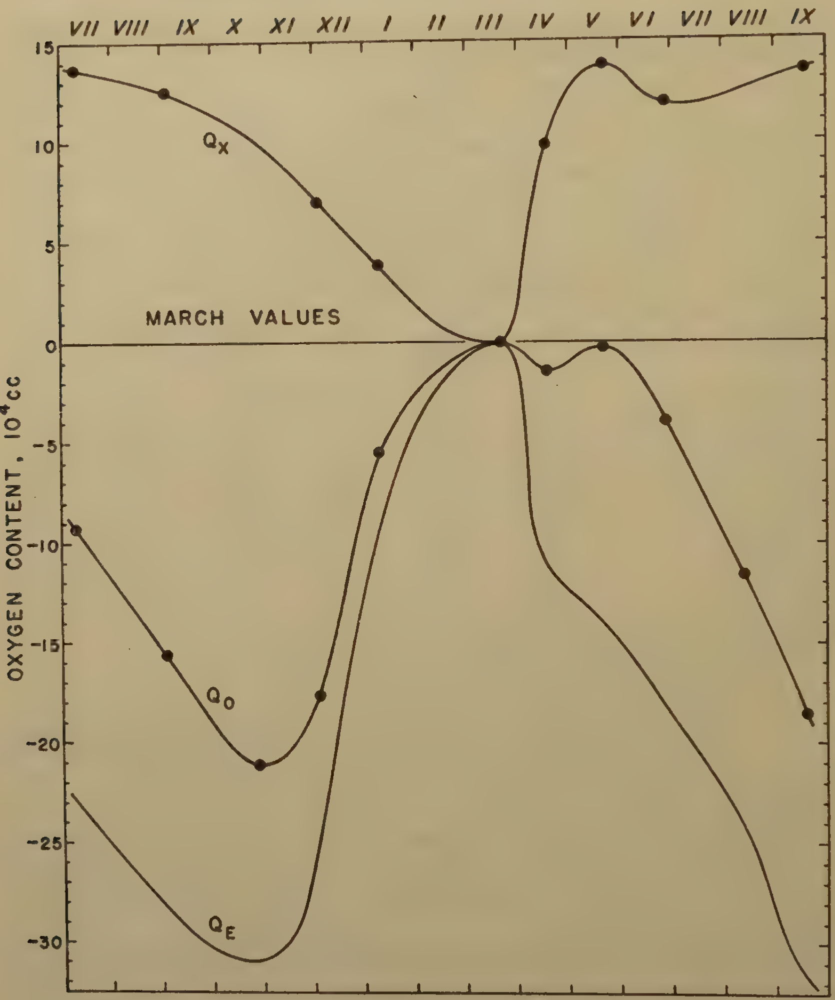


FIGURE 3. Seasonal values of ΔQ_o, the observed change in oxygen content, and ΔQ_x, the estimated change in oxygen content due to biological activity in the water beneath one m² of sea surface in Gulf of Maine, relative to the content observed in March. The curve ΔQ_E represents the quantities of oxygen estimated to have crossed one m² of sea surface between March and the time of observation. Negative values indicate losses of oxygen from the water.

Of the annual exchange across the sea surface only two-fifths may be attributed to the net production or consumption of oxygen by organisms. The remainder presumably results from the effects of temperature changes on the solubility of oxygen in sea water.

Metabolic processes and solubility effects due to temperature supplement one another in producing the surface exchange. However, their relative importance differs with the season. In April and May practically all surface exchange is accountable by oxygen produced in photosynthesis. From June through September the metabolic processes are balanced and the continued surface exchange of oxygen accompanies the warming of the water. During the winter both processes contribute to maintaining the exchange; two-fifths of the oxygen absorbed from the atmosphere is consumed by organic activity and three-fifths held in solution by the cooling water.

THE OXYGEN PRESSURE HEAD ACROSS THE SEA SURFACE

Oxygen may be presumed to move across the sea surface as a result of the difference between the partial pressure of oxygen in the atmosphere and the tension, or escaping tendency, of the dissolved oxygen in the surface water.

If it is assumed that the air in contact with the sea surface is at the temperature of the water and is saturated with water vapor, the oxygen pressure of the atmosphere, P, will be $0.21\ (1 - w)$ atmospheres where w is the tension of aqueous vapor. The oxygen tension of the water, p, is equal to the oxygen pressure of the atmosphere when the water is in equilibrium with the atmosphere or is "saturated." Under other conditions it follows from Henry's law that $p = P\ C_x/C_o$, where C_o is the oxygen content of sea water in equilibrium with one atmosphere of moist air and C_x is the oxygen content of the sample in question. The head of oxygen pressure across the sea surface is consequently given by

$$P - p = P(C_o - C_x)/C_o, \qquad (3)$$

when C_x is the oxygen content of water in the sea surface.

In order to evaluate $P - p$, data from the samples of water collected at a depth of one meter during each cruise have been averaged and the results are entered in Table II. In the direct estimation of P for application to the right-hand member of equation (3), allowance for the tension of water vapor in the atmosphere is inconsequential, since it reduces the value of P by only one or two per cent. In the estimation of $C_o - C_x$, however, it is important to take account of the water vapor in determining the value of C_o. Its value is proportional to P and will be reduced correspondingly as the vapor tension increases,

TABLE II. Estimates of $P - p$, the Pressure Head of Oxygen across Surface, and Averaged Data for Water at 1 Meter Depth on Which Estimate Is Based

Period of Observation	*Temp.* °C	*Salinity* ‰	*P* *atm*	C_o *cc/l*	C_x *cc/l*	$P - p$ *atm*
July 6–10, '33	15.18	31.96	0.206	5.83	6.12	−0.010
Sept. 4, '33	19.55	31.56	0.205	5.38	5.72	−0.013
Oct. 28–29, '33	11.02	32.18	0.207	6.32	6.26	0.002
Dec. 2–4, '33	7.11	32.86	0.208	6.84	6.65	0.006
Jan. 9, '34	5.24	32.96	0.208	7.14	6.69	0.103
March 22, '34	2.92	33.07	0.208	7.54	7.47	0.002
April 18, '34	4.72	32.75	0.208	7.24	7.65	−0.012
May 22, '34	9.19	31.65	0.208	6.59	7.27	−0.021
June 26–July 1, '34	14.95	31.93	0.207	5.89	6.28	−0.014
August 11, '34	18.45	31.96	0.205	5.49	5.89	−0.015
Sept. 19, '34	17.65	32.12	0.206	5.56	5.88	−0.012

the reduction amounting to as much as 2% at the warmer seasons. Since the difference between C_o and C_x does not exceed 10% and at most seasons is 6% or less, neglect of the correction for vapor tension will lead to substantial error in estimating $C_o - C_x$.

To determine C_o, values in the table of Fox (1909) for the quantity of oxygen dissolved in sea water in equilibrium with *dry* air have been corrected to take account of the reduction in partial pressure of oxygen in an atmosphere saturated with water vapor. An alignment chart constructed from these values has been used to estimate C_o from the observed temperature and salinity of the water samples for which C_x was measured.

The estimate of the values $P - p$ for the period of survey is plotted in Fig. 4. During the winter the pressure is positive, favoring a movement of oxygen from the atmosphere into the water. During the summer the pressure is reversed. The maximum positive value of 0.013 atmospheres was found in January, while the maximum negative value of 0.021 atmospheres occurred in May. The changes in direction come in October and March. These relations are consistent with the movements of oxygen across the sea surface deduced from measurements of oxygen and phosphate in the water column beneath the surface.

THE COEFFICIENT OF OXYGEN EXCHANGE

The rate of exchange of oxygen across the sea surface is defined theoretically by the equation

$$dQ/dt = E \cdot S \cdot (P - p), \tag{4}$$

where dQ/dt is the rate at which exygen enters the surface, S is the surface area and E is a constant, the exchange coefficient. E may conveniently be given the dimensions cc O_2/month × m² × atmosphere.

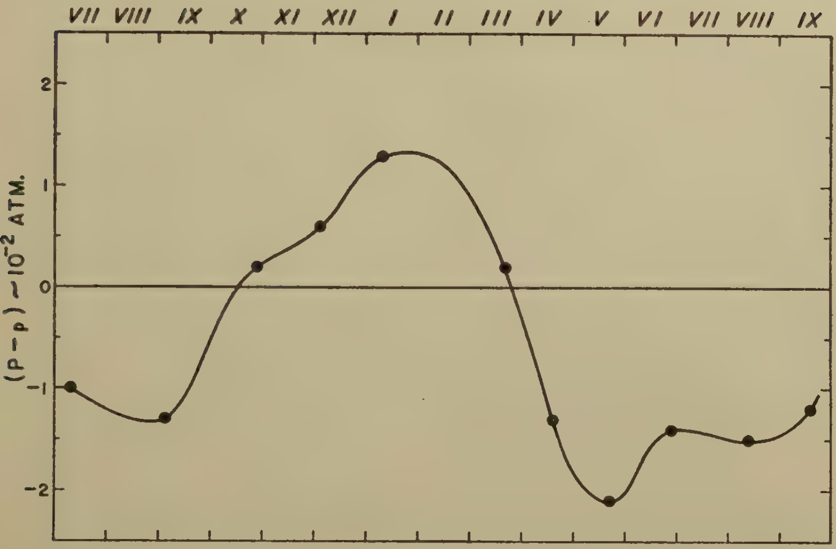


Figure 4. Seasonal variation in the pressure head of oxygen, $P - p$, across the surface of the Gulf of Maine. Negative values represent a pressure of oxygen in the water greater than in the atmosphere, a condition favoring the escape of oxygen from the sea water to the atmosphere.

Values of E may be estimated from equation (4), taking dQ/dt to equal the amount of oxygen entering the surface between successive times of observation divided by the time in months between the observations, and by expressing $P - p$ by the average of its values at the beginning and end of the intervening period. These approximations will not lead to large errors for periods when ΔQ_E and $P - p$ are relatively constant or are changing at nearly constant rates. The estimations are not reliable when the curve describing these functions are undergoing inflection, since at such times the average values may differ greatly from the properly weighted mean. Estimates of the value of E obtained in this way are given in Table III.

The three estimates made between the end of October and the latter part of March when $P - p$ is positive, indicating a movement of oxygen into the sea water, yield an average value for E of 11.3×10^6. The value for January 9–March 22 is suspect, because the average value of $P - p$ is unreliable, being obtained when the curve is undergoing inflection. The two values obtained between October 29 and January 9 are in close agreement and give the most reliable estimate

TABLE III. Estimates of the Coefficient for the Exchange of Oxygen across the Surface of the Gulf of Maine

	1	*2*	*3*	*4*	*5*
Period of Observation	Δt	ΔQ	$\Delta Q/\Delta t$	$(P - p)_m$	E
July 8–Sept. 4	1.8	−5.2	−2.9	−0.0155	2.5
Sept. 4–Oct. 29	1.8	−2.9	−1.6	−0.0055	2.9
Oct. 29–Dec. 3	1.2	6.4	5.3	0.0040	13.2
Dec. 3–Jan. 9	1.2	15.4	12.8	0.0095	13.5
Jan. 9–Mar. 22	2.4	9.3	3.9	0.0075	5.2
Mar. 22–Apr. 18	0.9	−11.2	−12.4	−0.0050	24.8
Apr. 18–May 22	1.1	−3.8	−3.5	−0.0165	2.1
May 22–June 28	1.1	−4.3	−3.9	−0.0175	2.2
June 28–Aug. 11	1.5	−5.9	−3.9	−0.0145	2.7
Aug. 11–Sept. 19	1.3	−7.8	−6.0	−0.0135	4.4

Column *1* Δt is interval between observations—in months.
2 ΔQ is quantity of oxygen entering surface during interval, *i. e.*, the change in ΔQ_E values of Table II. Unit: 10^4 cc/m².
3 $\Delta Q/\Delta t$ is mean rate of exchange during interval. Unit: 10^4 cc/m² × months.
4 $(P - p)_m$ is mean of values of oxygen pressure head $P - p$ at beginning and end of interval. Unit: atmospheres.
5 E is exchange coefficient of oxygen. Unit: 10^6 cc/month × m² × atmosphere.

for E for the period when oxygen is entering the sea surface as 13×10^6.

Of the estimates made during the remainder of the year, when $P - p$ is negative, that for March 22–April 18 is unusually large. Its reliability may be questioned, because at that time the estimate of the oxygen exchange is based almost wholly on changes in the phosphate concentration of the upper layers of water. The resulting value for E is entirely subject to the errors in this estimation. The six remaining estimates are in fair agreement and yield an average value for E of 2.8×10^6. The exchange coefficient for the period when oxygen is leaving the sea may be fixed approximately at this value.

These estimations indicate that the exchange coefficient of oxygen at the sea surface is greater during the winter period when oxygen is being taken up by the water than during the summer period when the movement of oxygen is in the opposite direction. The values for E for these two periods appear to be approximately 13×10^6 and 2.8×10^6 respectively.

DISCUSSION OF THE COEFFICIENT OF OXYGEN EXCHANGE

The only pertinent data on the rate of movement of oxygen across an air-water interface of which I am aware are those of Bohr, Krogh, and Adeney and Becker. In studying the absorption of oxygen by small bubbles at 37° C, Krogh (1910) obtained a value of 0.0776 cc per minute per cm² per atmosphere, which, when converted to the dimensions in which E has been expressed, equals 33.5×10^6 cc per month per m^2 per atmosphere. Adeney and Becker (1920) measured the quantity of oxygen leaving a relatively large bubble in its passage through unsaturated water. Estimates by Dorsey (1940: 552–555) based on their data give values of about 0.017 cc per minute per cm^2 per atmosphere or, when converted into the presently employed units, 7.3×10^6 cc per month per m^2 per atmosphere. Adeney and Becker's results indicate that the coefficient is about 5% lower for sea water and that it is not materially changed by temperature.

The present estimates for the sea surface agree in magnitude with these measurements. The values of Adeney and Becker are actually spanned by the estimates made for the summer and winter periods.

The difference in the value of the exchange coefficient between summer and winter is sufficiently distinct to suggest that some real differences in the conditions near the sea surface at these seasons affect the exchange. Before accepting this conclusion it is desirable to examine the assumption that the air immediately over the sea surface is saturated with water vapor. The introduction of the correction for water vapor causes the values of $P - p$ to be smaller in winter and larger in summer than would be the case if the atmosphere were assumed to be dry or incompletely saturated. The values of E are consequently made larger in winter and smaller in summer. Thus the application of this correction might produce seasonal differences in the estimated value of E of the sort observed.

To determine whether the assumed saturation of the atmosphere could produce the seasonal difference obtained for E its values for the periods December 2–January 9 and June 26–August 11 have been calculated, assuming the atmosphere over the sea surface to be dry. The results give values of 10.5×10^6 for the former period and $3.7 \text{ x } 10^6$ for the latter. The results obtained, assuming a saturated atmosphere, were 13.5×10^6 and 2.7×10^6 respectively. It is evident that the difference in the winter and summer values of E cannot be attributed in any major degree to possible error in this assumption.

The possibility that the difference in the value of E depends directly on the seasonal difference in temperature is ruled out by Adeney and Becker's measurements on bubbles.

Two conditions suggest themselves in explanation of the seasonal difference in E. The values are calculated on the assumption of a flat sea surface of unit area. Waves, spindrift, and the bubbles entrained by white caps, which all tend to increase the air-water interface within a given area, are much more prominent in winter, favoring higher values of E at that time. Also, conditions near the sea surface in summer are much more stable than in winter. There is greater probability that the air immediately over the surface is saturated with water vapor and that small scale gradients of oxygen concentration exist immediately beneath the sea surface. The use of oxygen analyses from samples collected at a depth of one meter in estimating the pressure of oxygen in the surface water leaves much to be desired. In winter, when unstable conditions favor boisterous mixing, this procedure is probably adequate, but in summer it is likely that the water at the sea surface is more nearly in equilibrium with the atmospheric oxygen than the sample collected. This appears to be the most important consideration in accounting for the seasonal difference in the estimated values of the coefficient of oxygen exchange. Taken together the several factors considered are sufficient to account for the results.

SUMMARY

1. It is estimated that about 30×10^4 cc of oxygen per m^2 leaves the surface of the Gulf of Maine between the end of October and the latter part of March, a like amount re-entering from the atmosphere during the alternating period.

2. The net production and consumption of oxygen by organisms is sufficient to account for two-fifths of this exchange, the remainder being attributable to the effects of temperature on the solubility of oxygen.

3. In April and May the production of oxygen in photosynthesis is sufficient to account for the entire surface exchange; during the summer the exchange is attributable chiefly to decreasing solubility resulting from the warming of the water. In late fall and winter, excess oxygen consumption and increasing solubility are both responsible for the exchange.

4. The pressure head of oxygen across the sea surface has been determined and found to agree in direction with the estimated movement of oxygen.

5. The exchange coefficient of oxygen is estimated to be about 13×10^6 cc per m^2 per atmosphere per month in winter. In summer a value of 2.8×10^6 is obtained.

6. The seasonal difference in the values obtained is attributed to conditions at the sea surface, such as sea state and stability, which cause errors in the measurements and estimations.

REFERENCES

Adeney, W. E. and H. G. Becker

1920. The determination of the rate of solution of atmospheric nitrogen and oxygen by water. Part II. Sci. Proc. R. Dublin Soc., *15:* 609–628.

Bigelow, H. B.

1927. Physical oceanography of the Gulf of Maine. Bull. U. S. Bur. Fish., *40* (2): 510–1027.

Cooper, L. H. N.

1938. Redefinition of the anomaly of the nitrate-phosphate ratio. J. Mar. biol. Ass. U. K., N. S. *23:* 179.

Dorsey, N. E.

1940. Properties of ordinary water-substance. Reinhold, New York. 673 pp.

Edmondson, W. T. and Y. H. Edmondson

1947. Measurements of production in fertilized saltwater. J. Mar. Res., *6:* 228–246.

Fleming, R. H.

1941. Composition of plankton and units for reporting populations and production. Proc. 6th Pacific Sci. Congr., *3:* 535–540.

Fox, C. J. J.

1909. On the coefficients of absorption of nitrogen and oxygen in distilled water and sea water, and of atmospheric carbonic acid in sea water. Trans. Faraday Soc., *5:* 68–87.

Krogh, August

1910. Some experiments on the invasion of oxygen and carbonic-oxide into water. Skand. Arch. Physiol., *23:* 224–235.

Redfield, A. C.

1934. On the proportions of organic derivatives in sea water and their relation to the composition of plankton. James Johnstone Memor. Vol.: 176–192. University Press, Liverpool. 348 pp.

1942. The processes determining the concentration of oxygen, phosphate and other organic derivatives within the depths of the Atlantic Ocean. Pap. phys. Oceanogr. Meteorol., *9* (2): 1–22.

Redfield, A. C., H. P. Smith, and B. H. Ketchum

1937. The cycle of organic phosphorus in the Gulf of Maine. Biol. Bull. Woods Hole, *73:* 421–443.

SOME SEASONAL CHEMICAL CHANGES IN THE OPEN OCEAN[1]

By

NORRIS W. RAKESTRAW AND DAYTON E. CARRITT

Scripps Institution of Oceanography
University of California
La Jolla, California

Several studies have been made of the seasonal changes in the chemical composition of sea water, especially of the so-called "nutrient substances." But most of these have been in localities relatively near shore and in the temperate zones where climatic fluctuations are relatively great. During the years 1937 and 1938 an opportunity arose to study such changes in the open ocean in connection with an investigation of the variations of the Gulf Stream (Iselin, 1940).

Three stations were chosen on a line between Montauk Point, N. Y., and Bermuda as follows: Station A, on the Continental Shelf about 30 mi. SE. of Montauk Point (Lat. 40° 40′ N; Long. 71° 45′ W); Station B, in the slope water northwest of the current of the Gulf Stream (Lat. 39° 00′ N; Long. 70° 40′ W); Station C, in the Sargasso Sea southeast of the Gulf Stream (Lat. 35° 30′ N; Long. 67° 30′ W). These stations were visited seven times during the year from October 1937 to September 1938, and determinations were made of the following chemical properties and substances: pH, alkalinity, phosphate, nitrate, nitrite, ammonia. These, along with temperature and salinity, would have afforded a comprehensive chemical picture of conditions, but, unfortunately, it was seldom possible to complete such an extensive chemical program at any one station. There are many gaps, therefore, in the chemical data, but enough information was obtained to suggest some conclusions.

The methods used were those which have been described before: pH by the indicator method using cresol red and the Hastings bicolorimeter without buffers (Mitchell and Taylor, 1935; Mitchell, *et al.*, 1936); alkalinity by the method of Mitchell and Rakestraw (1933); phosphate, nitrate, and nitrite by the standard methods of Atkins (1923), Harvey (1926), and Griess (1879) and Ilosvay (1889), respectively; ammonia by a modification of the method of Krogh (1934).

---

[1] Contribution No. 452 from the Woods Hole Oceanographic Institution.

The concentration of free carbon dioxide was calculated from the pH, alkalinity, temperature, and salinity, by the procedure of Buch, *et al.* (1932).

The times at which the stations were visited were as follows:

Cruise I, October 2–9, 1937
II, December 6–16, 1937
III, January 3–5, 1938
IV, April 6–12, 1938
V, May 29–June 5, 1938
VI, July 10–17, 1938
VII, August 24–September 1, 1938

The significant results are shown in Figs. 1–6.

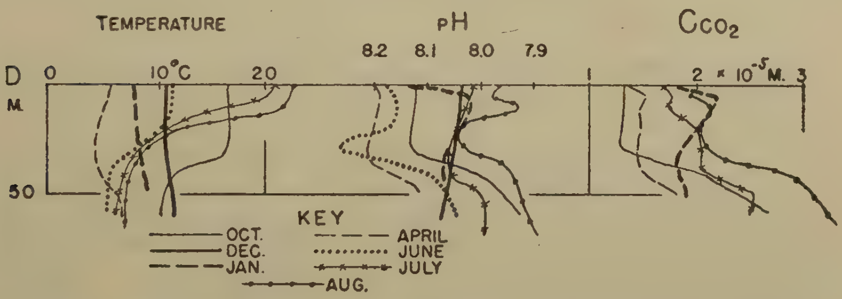


Figure 1. Temperature, pH, and free carbon dioxide concentration (C_{CO_2}) at Station A. C_{CO_2} in moles × 10^{-5}.

From Figs. 1 and 2 it may be seen that the conditions at Station A were substantially the same as at other similar locations which have been studied heretofore, as, for example, the English Channel and the Gulf of Maine. During the summer months, when a strong thermocline develops, the upper levels become depleted of nutrient substances. In the winter a temperature inversion takes place, the whole water column becomes homogeneous or even unstable, and vertical circulation brings about replenishment from the richer deep water. The curves for pH and carbon dioxide concentration show vigorous biological activity going on in summer, with less in winter.

This is one of the few occasions on which ammonia and nitrite determinations have been made simultaneously, and it is important to notice that the two follow the same general pattern. Both are found at the surface during the winter and both are generally lacking there in the summer. This lends some support to the view, developed from

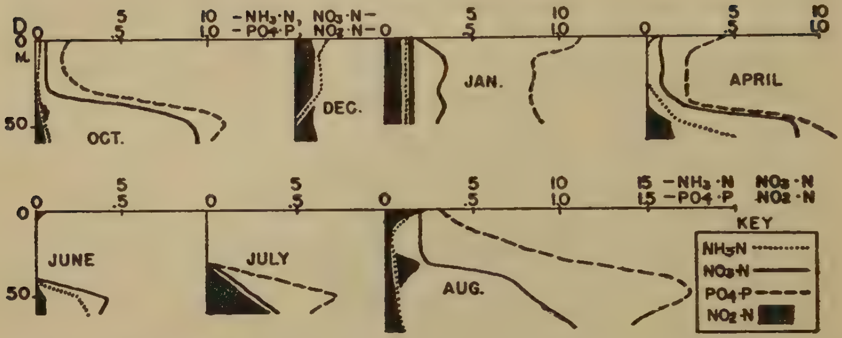


Figure 2. Seasonal distribution with depth at Station A of: ammonia-nitrogen, nitrate-nitrogen, phosphate-phosphorus, nitrite-nitrogen, in microgram atoms per liter. The upper scale of units is for ammonia and nitrate, the lower for phosphate and nitrite.

experimental studies (von Brand, *et al.*, 1937), that ammonia, which is the first product of organic decomposition, is the precursor of nitrite in the sea.

Fig. 3 shows the relation between temperature, oxygen, pH, and carbon dioxide at Station B. It will be noted that the position of the oxygen minimum follows the thermocline up and down with the seasons and, as usual, is slightly above the pH minimum. The carbon dioxide concentration, more clearly than the pH, also follows the fluctuations of the thermocline.

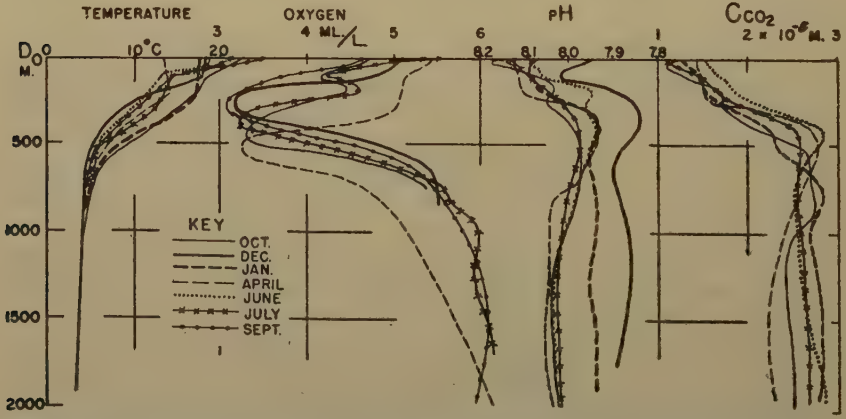


Figure 3. Temperature, oxygen, pH, and free carbon dioxide concentration at Station B. Oxygen in ml/l., carbon dioxide in moles $\times 10^{-5}$.

The curves for pH at different times of the year do not coincide as closely as might be expected. There is little reason to believe that there should be as great differences at, say 1,000 meters, as are indicated. In the method for this determination, which makes use of the dissociation constant of the indicator rather than a set of standard buffers, significant errors may be introduced by using different batches of indicator, or even the same indicator over a considerable period of time. The glass electrode had not at that time been developed to its present state of reliability and convenience, otherwise the indicator could have been repeatedly standardized by this means. Consequently, although the relative values on each cruise are probably reliable, it is not safe to compare too closely the results obtained on different cruises. It would have been possible to assume that pH changes are insignificant below the thermocline and to have shifted all the curves to the right or left to make them pass through a common point at some standard depth (1,000 or 1,500 meters). The validity

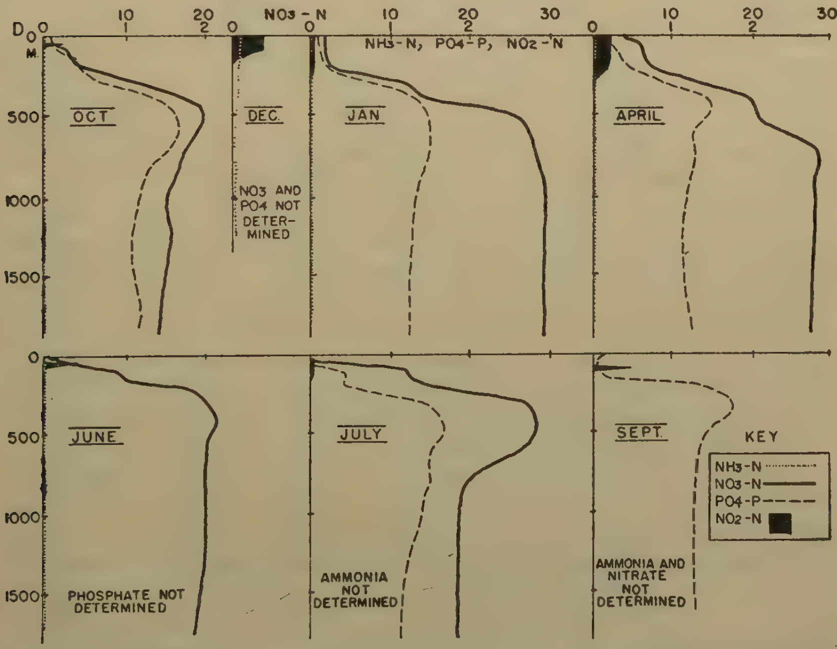


Figure 4. Seasonal distribution with depth at Station B of: ammonia-nitrogen, nitrate-nitrogen, phosphate-phosphorus, nitrite-nitrogen, in μg. at./l. The upper scale of units is for nitrate, the lower for ammonia, phosphate, and nitrite.

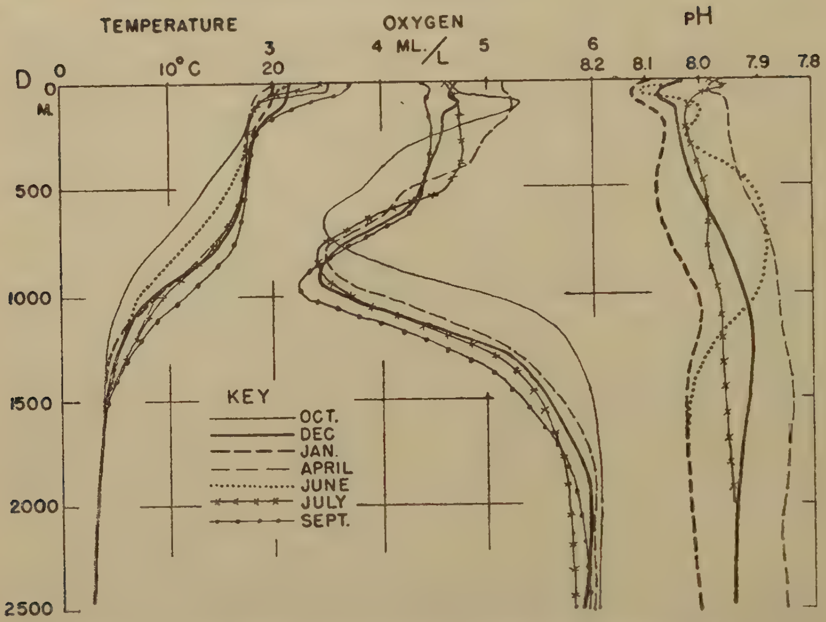


Figure 5. Temperature, oxygen, and pH at Station C. Oxygen in ml./l.

of this assumption is not certain, however, and the results of such an adjustment do not bring out any clearer relations.

The pH value is involved in the calculation of the carbon dioxide concentration, and since the variations in the latter (for five cruises) agree well with the variations in oxygen, it would seem that the pH values on at least these five cruises are fairly reliable.

Fig. 4, showing the distribution of nitrate, nitrite, ammonia, and phosphate at Station B, reveals two interesting facts. First, the positions of the phosphate and nitrate maxima fluctuate with the thermocline, which might have been expected from their close relation to the oxygen minimum. Secondly, in the upper levels, the same changes take place—although somewhat less clearly—as at Station A.

The temperature is lower at the surface in winter, and the upper layer becomes homogeneous. Vertical circulation brings nitrate and phosphate up from the transition layer to the surface. In summer, stable stratification again sets in and nutrients disappear by consumption. Nitrite also appears at the surface in winter and disappears in summer. In this case ammonia did not show as clear a relation to nitrite as at the shallower station; only in December was it found in small amounts, principally in the upper levels.

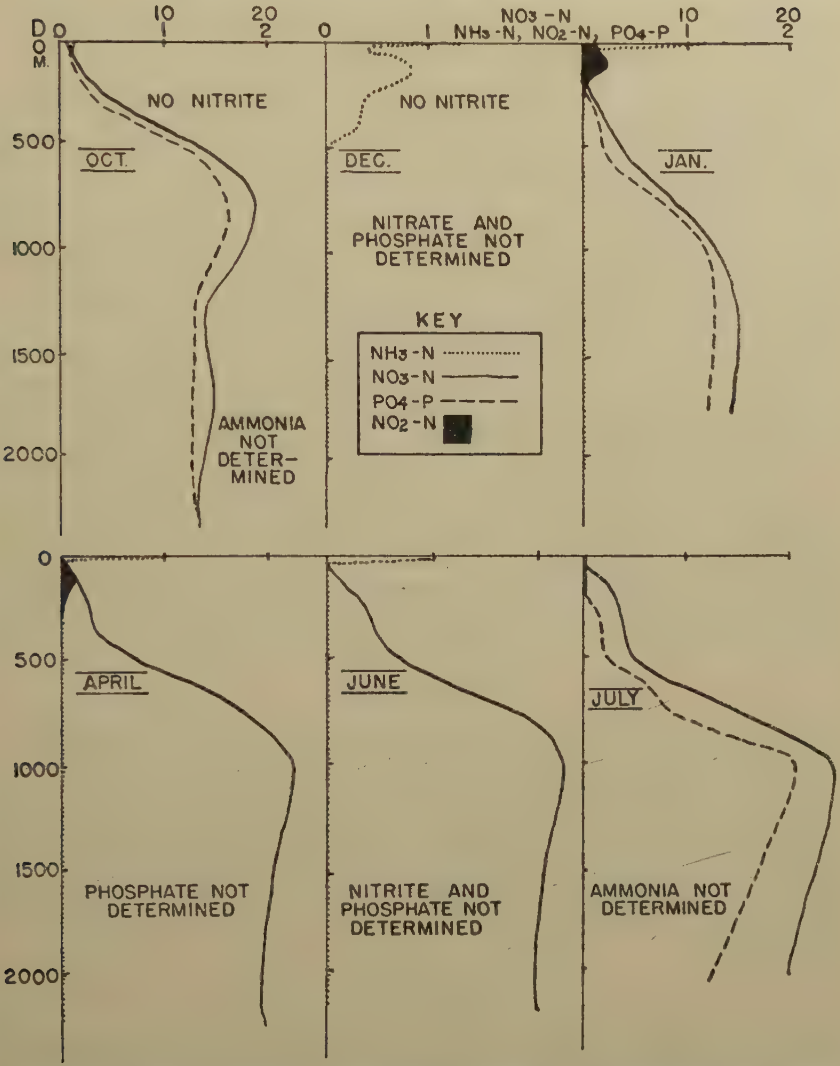


Figure 6. Seasonal distribution with depth at Station C of: ammonia-nitrogen, nitrate-nitrogen, phosphate-phosphorus, nitrite-nitrogen, in μg. at./l. The upper scale of units is for nitrate, the lower for ammonia, phosphate, and nitrite.

Nitrite appeared in the surface in December and remained through January and April; in June, July, and September, however, it had disappeared from the surface. Somewhat lower down, in a narrow zone extending to 100 or 200 meters, significant amounts were still to be found. This is in accord with a previous study (Rakestraw, 1936).

The conditions at Station C were in general the same as those at Station B, although in some respects less clear. Fig. 5 shows the same relation between oxygen distribution and the temperature structure. The remarks already made about the pH results are equally true at this station. Although the surface values for nutrients are not conclusive, partly because of incomplete data from the December cruise, there is nevertheless a strong indication that there is not as much replenishment from the transition zone as at Station B. Thermal stratification is not completely destroyed in the winter.

For some reason, significant amounts of ammonia seem to have been present commonly in the surface, although only in December did these extend further down. Notwithstanding this, no nitrite was present at this time. This is different from the condition at Station B but it may be highly significant. It suggests that nitrite, which is never abundant in this semitropical oceanic water, had not yet been formed by oxidation of its precursor, ammonia. The ammonia, on the other hand, may have been a recent result of organic decomposition. By January, however, the oxidation had taken place, and the nitrite persisted through April until June and July, when only a small trace was found.

SUMMARY

At least to the extent to which seasonal changes in temperature, etc., take place in the upper layers in the open ocean, the seasonal variations in nutrients and other chemical factors are the same as those which occur in shallow inshore regions of the temperate zones. In addition, however, the distribution of chemical constituents and properties varies with the movements of the water mass and the changes in its physical properties, such as the rise and fall of the thermocline.

REFERENCES

Atkins, W. R. G.
1923. The phosphate content of fresh and salt waters in relationship to the growth of the algal plankton. J. Mar. biol. Ass. U. K., N. S. *13:* 119–150.

von Brand, T., N. W. Rakestraw and C. E. Renn
1937. The experimental decomposition and regeneration of nitrogenous organic matter in sea water. Biol. Bull. Woods Hole, *72:* 165–175.

Buch, K., H. W. Harvey, H. Wattenberg and S. Gripenberg
1932. Über das Kohlensäuresystem in Meerwasser. Rapp. Cons. Explor. Mer, *79:* 1–80.

Griess, Peter
1879. Bemerkungen zu der Abhandlung der H. H. Weselsky und Benedikt "Ueber einige Azoverbindungen." Ber. dtsch. chem. Ges., *12:* 426–428.
Harvey, H. W.
1926. Nitrate in the sea. J. Mar. biol. Ass. U. K., N. S. *14* (1): 71–88.
Ilosvay, L.
1889. L'acide azoteux dans la salive et dans l'air exhalé (VI). Bull. soc. chim. Fr., (3) *2:* 388–391.
Iselin, C. O'D.
1940. Preliminary report on long-period variations in the transport of the Gulf Stream system. Pap. phys. Oceanogr. Meteorol., *8* (1): 1–40.
Krogh, A.
1934. A method for the determination of ammonia in water and air. Biol. Bull. Woods Hole, *67:* 126–131.
Mitchell, P. H., K. Buch and N. W. Rakestraw
1936. The effect of salinity and temperature upon the dissociation of cresol red and phenol red in sea water. J. Cons. int. Explor. Mer, *11* (3): 183–190.
Mitchell, P. H. and N. W. Rakestraw
1933. The buffer capacity of sea water. Biol. Bull. Woods Hole, *65:* 437–451.
Mitchell, P. H. and I. R. Taylor
1935. The dissociation constant of cresol red in sea water. J. Cons. int. Explor. Mer, *10* (2): 169–172.
Rakestraw, N. W.
1936. The occurrence and significance of nitrite in the sea. Biol. Bull. Woods Hole, *71:* 133–167.

THE SEA SAMPLER[1]

By

ATHELSTAN F. SPILHAUS AND ARTHUR R. MILLER

Woods Hole Oceanographic Institution[2]
Woods Hole, Massachusetts

INTRODUCTION

The original sea sampler, as described by Spilhaus (1940), consisted of six bottles, each built somewhat like a small version of a standard Nansen bottle. These surrounded a cylinder containing a releasing block which was operated by an ordinary bathythermograph pressure element. The tripping block, which was moved by the pressure element, was scored and grooved to accommodate a series of triggers, one for each bottle, which closed the stop-cocks in turn as the whole instrument reached predetermined pressures. The tripping block was arranged so that the triggering points of the different bottles could be varied at will. The instrument was designed for use with the bathythermograph, which was described earlier (Spilhaus, 1938 and 1940), and therefore the next step was to combine the two instruments so that a single pressure assembly operated the tripping block for the sea sampler and carried the slides for the record of the bathythermograph temperature.

Extensive observations with the combined thermographic sea sampler were made during the summer of 1940 and are described by Spilhaus (1941). This series of observations, designed to demonstrate the rapid measuring technique made possible by this instrument, provided 200 soundings, giving 200 temperature-depth curves and 1,400 salinity and oxygen values, all obtained in a continuous period of three and a half days. This represents an average of one sounding (giving temperature against depth, and seven oxygen and salinity values) every twenty-two and one-half minutes, or every two miles at the speed of the vessel utilized, which was 5½ knots. During World War II, and because of its immediate application to the operating problems of underwater sound, considerable refinement of the bathythermograph was made by Dr. Ewing and his associates at the Woods

---

[1] Contribution No. 440 from the Woods Hole Oceanographic Institution.

[2] The construction and testing was done in the course of work by the Woods Hole Oceanographic Institution under contract with the Office of Naval Research.

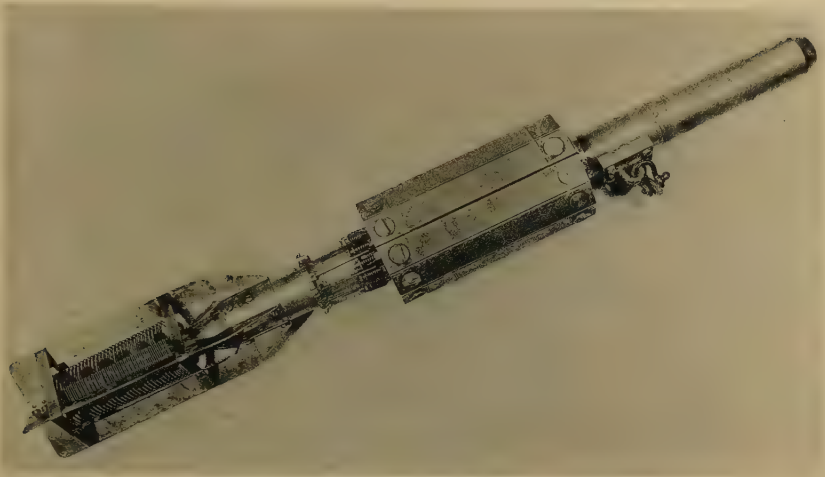


Figure 1. The new thermographic sea sampler.

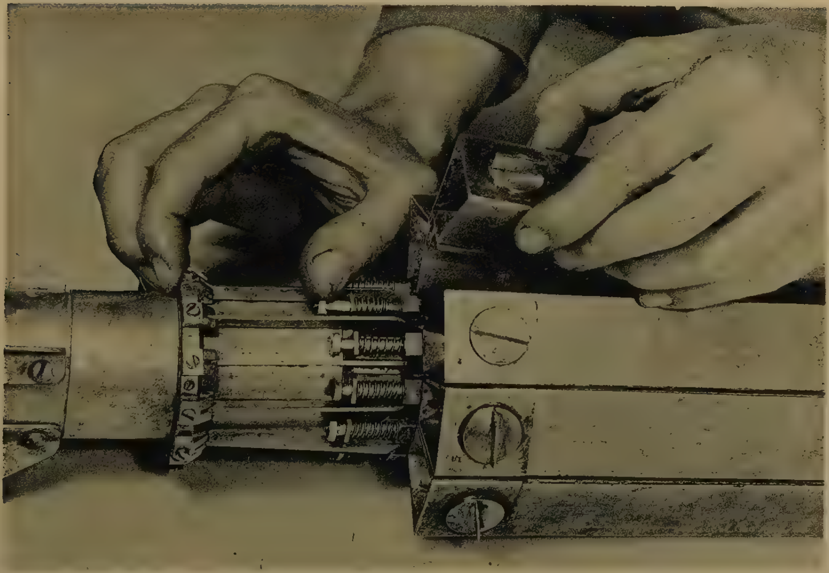


Figure 2. The sampler bottles showing method of removal.

Hole Oceanographic Institution, but improvement of the sea sampler was deferred until last year when the work described below was started.

DESCRIPTION OF THE NEW THERMOGRAPHIC SEA SAMPLER

In principle, the present instrument differs little from the thermographic sea sampler described earlier. However, two important improvements have been made. These are, first, streamlining the bottles so that the whole instrument may be used with the vessel underway at higher speeds, and second, arranging the tripping mechanism so that the samples are taken on the way up from the lowest depth reached rather than on the way down, as was the case in the earlier instrument. The number of bottles is increased to twelve, and the whole design is so arranged that in one or two minutes all the bottles and triggers can be removed without the use of any tools; the instrument is then identical to an ordinary bathythermograph and may be used as such.

STREAMLINING

So that this addition to the bathythermograph would give as little drag as possible, the bottles, twelve in number, were made in a trapezoidal shape and were fitted around the instrument to form a smooth cylinder of not much greater diameter than the width of the fins on the bathythermograph (Figs. 1, 2 and 3). Furthermore, to maintain the smooth exterior of the cylinder, the connecting bar and spring, which are necessary on each bottle to effect the simultaneous closing of the two rotary stop-cocks, were placed on the inside surface of the bottle; that is, on the side of the bottle that faces the bathythermograph

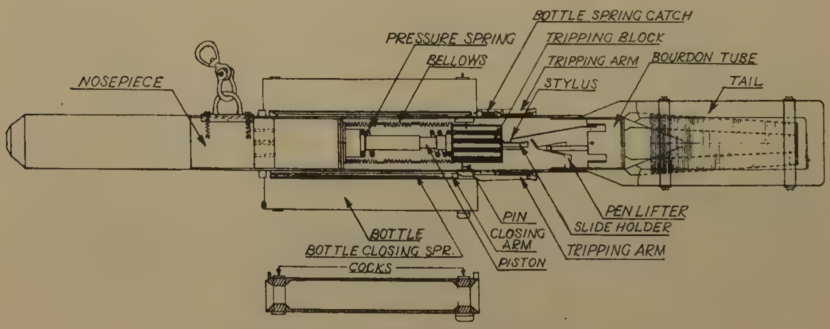

BOTTLE SPRING CATCH
PRESSURE SPRING
TRIPPING BLOCK
BELLOWS
TRIPPING ARM
BOURDON TUBE
NOSEPIECE
STYLUS
TAIL
PEN LIFTER
PIN
SLIDE HOLDER
CLOSING ARM
BOTTLE
TRIPPING ARM
BOTTLE CLOSING SPR.
PISTON
COCKS

Figure 3. Diagram of Sea Sampler

cylinder. This has the additional advantage that the mechanism is protected from damage. The result of this streamlining as it affects the rate of fall of the combined thermographic sea sampler is well illustrated in Fig. 4, which shows a comparison between the drop of a bathythermograph and that of the sea sampler (with six of the twelve bottles attached). These tests, which show that the sampler dropped as fast as the bathythermograph, were made at a speed of 8½ knots and therefore indicate that the former should be utilizable at as high speeds as the latter. In the tests made, it took longer to bring in the sea sampler than the bathythermograph, but this is merely a matter of the power of the winch.

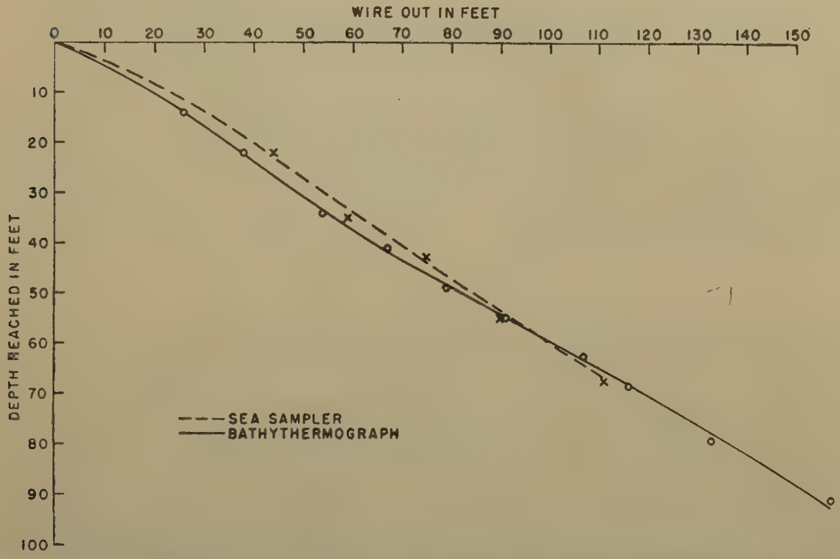


Figure 4. Comparison of the Rate of Drop of Sea Sampler and Bathythermograph with respect to Length of Wire Out while Underway at 8½ Knots

TRIPPING MECHANISM

The tripping mechanism is now designed so that water is trapped as the whole instrument ascends to the sea's surface. Thus the sample is never immersed in denser water, and more thorough washing of the bottle is accomplished. A further alteration is that the adjustment of the depth of tripping is now made by simple exchange of the tripping bars on the exterior of the instrument. In the earlier model these depths could be changed only by replaceable grooves in the tripping

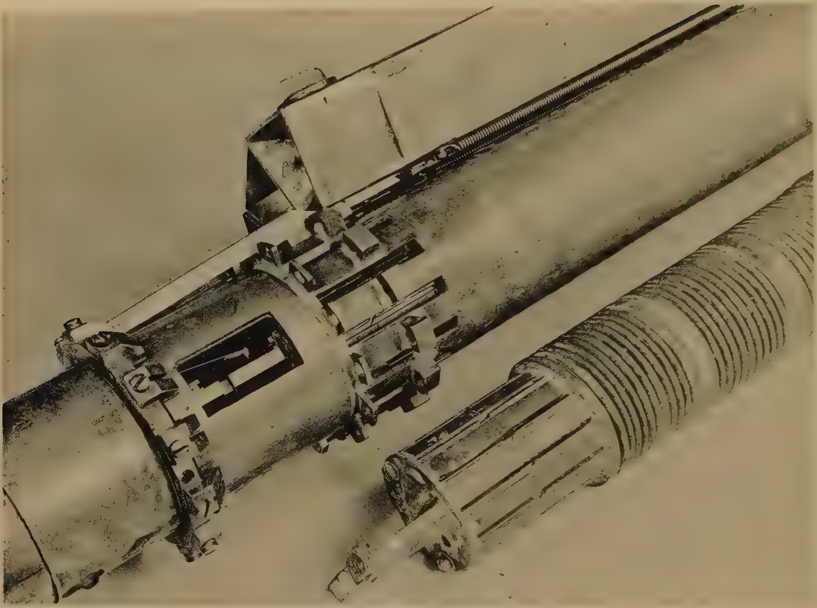


Figure 5. Tripping mechanism.

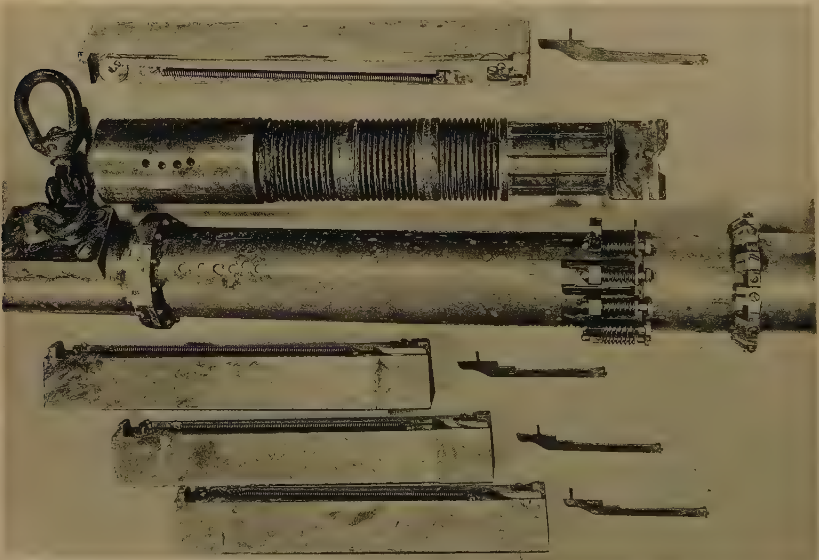


Figure 6. Sea sampler with bottles and triggers removed and a pressure unit carrying the tripping block.

block, which was inconvenient, since the pressure mechanism had to be withdrawn in order to do this. Figs. 5 and 6 and the diagram in Fig. 7 indicate the manner in which the tripping is achieved in the present instrument. Calibration of tripping depth can be carried out either in a pressure tank or by the use of an auxiliary device which withdraws the pressure element and then pushes it back into place, thus simulating on the bench the compression of the sylphon and the spring due to water pressure. Fig. 8 shows calibration curves obtained by both methods. In this case, in order to show up any frictional errors to their maximum extent, a shallow pressure unit (120 feet) was utilized. Owing to the size of the pressure tank available, only three bottles could be used at one time in the tank tests. The method of test was to place the sea sampler in the tank, then build up the pressure until the deepest bottle was in a cocked position. The water pressure was then slowly decreased until a sharp crack, indicating the tripping point of the bottle, was heard. The pressure at this point was noted for each bottle and with each tripping bar. Repeated tests showed excellent consistency, any slight variation being attributable to the accuracy of reading the pressure scale. The bench tests were conducted in a similar manner, but, by withdrawing the pressure assembly and then pushing it back with the device described above. At the point of tripping, the stylus of the thermom-

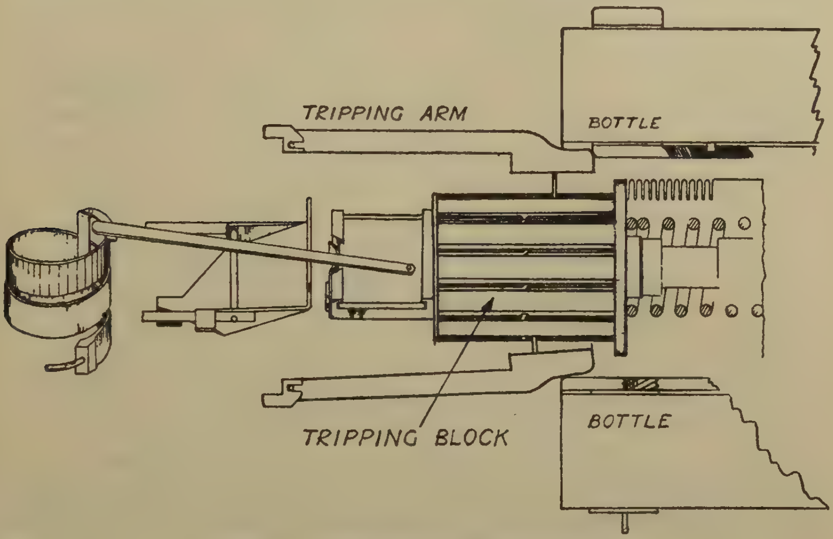


Figure 7. Diagram of Tripping Mechanism

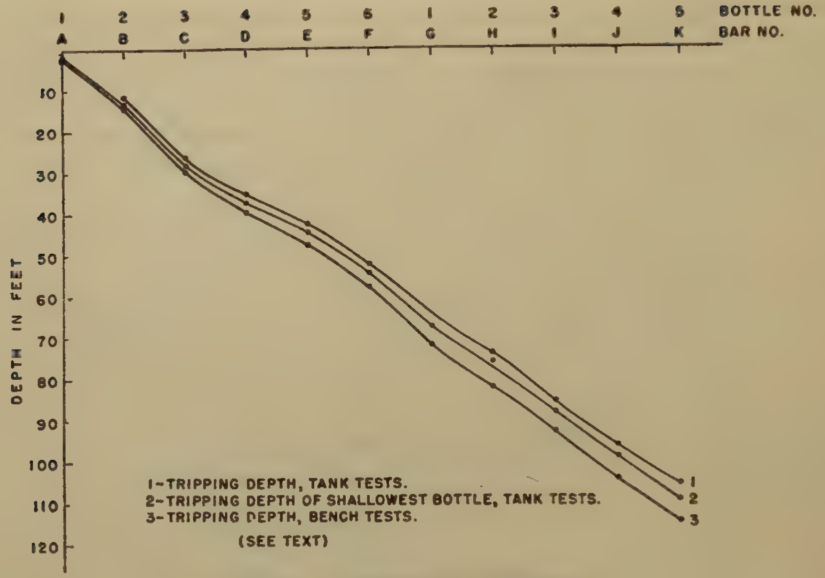


Figure 8. Calibration of Sea Sampler

eter was moved slightly on the smoked slide to mark the depth. It may be noted that the shallowest bottle in the group will trip two feet below its normal calibration. This is believed to be a frictional effect which may be minimized when all twelve bottles are in position around the instrument so that the sideways pressure of the tripping bars is evenly distributed rather than one-sided as in the case of these tests.

TRIALS AT SEA

Some tests of the thermographic sea sampler were made from the research vessel Balanus in order to compare salinities obtained by it with those obtained by the Nansen bottle technique. The purpose of these tests was twofold: first, to check the over-all performance of the sea sampler under conditions of practical use, and second, to study the effect of throttling, which prevents the complete interchange of water within the sampling bottles due to the constriction of the stop-cocks. The shallow bathythermograph pressure unit was used in order to magnify any frictional effect, and the station was chosen in the Narrows of New York Harbor, where an extreme salinity gradient was found. Under these conditions any errors due to the throttling of the stop-cocks of the sample bottles would be shown up by marked differences of salinity between the sea sampler and the corresponding Nansen bottle. Test numbers 1, 2 and 3 (Table IA) were made at this station in a depth of about 78 feet of water just off Quarantine. The salinities here were very low, ranging from about 9 ‰ at the surface to 27 ‰ at 50 feet. This gradient of salinity provided ideal and stringent conditions for tests on the throttling effect of the bottles. Numbers 1 and 2 were made by lowering Nansen bottles set at exactly the tripping depth of the six bottles used on the sea sampler. The

TABLE I. Salinities (‰) Obtained in Tests of the Sea Sampler and Nansen Bottles on Balanus Cruise B-19, March 25, 1948

A. Stations I, II and III at Anchor in Narrows, near Quarantine, New York (Lat. 40° 36.8′ N., Long. 74° 03.4′ W.)

Depth of Nansen Bottles and Tripping Points of Sea Sampler, m.	*I** *0950 hr.*		*II** *1025 hr.*		*III†* *1130 hr.*	
	Nansen	*Sea Sampler*	*Nansen*	*Sea Sampler*	*Nansen*	*Sea Sampler*
1.2	10.91	12.85	9.56	10.95	10.21	10.86
4.0	11.51	16.51	10.14	14.42	10.72	10.99
7.6	17.38	——	15.12	22.65	——	——
9.8	22.03	23.82	21.37	24.83	——	——
13.7	23.84	25.77	25.95	——	——	——
16.2	25.61	26.47	26.78	——	——	——
Depth (max.) on B. T. (ft.)		77		78	55	68

B. Stations IV and V at Anchor, Middle of Sandy Hook Channel, New York
Station VI Underway, Sandy Hook Channel
(Lat. 40° 28.7′ N., Long. 74° 01.2 W.)

	*IV** *1325 hr.*		*V** *1345 hr.*		*VI‡* *1410 hr.*	
1.2	14.51	14.87	14.63	15.12		——
4.0	15.17	15.81	14.97	15.81	None	14.76
7.6	15.82	16.04	15.86	16.06	Underway	15.91
9.8	16.09	——	15.97	16.11		——
13.7	16.20	——	16.02	17.79		——
16.2	16.94	——	17.07	——		——
Depth (max.) on B. T. (ft.)	55	65	55	67		70

B. T. 6 ft. below lowest Nansen Bottle. Vessel speed 8½ knots.

* Sea Sampler's vertical rate of ascent, 60 ft. in 10 sec.

† Sea Sampler brought up very slowly with stops.

‡ Sea Sampler's vertical rate of ascent, 60 ft. in 35 sec.

Temperature approximately 40° F. and homogeneous.

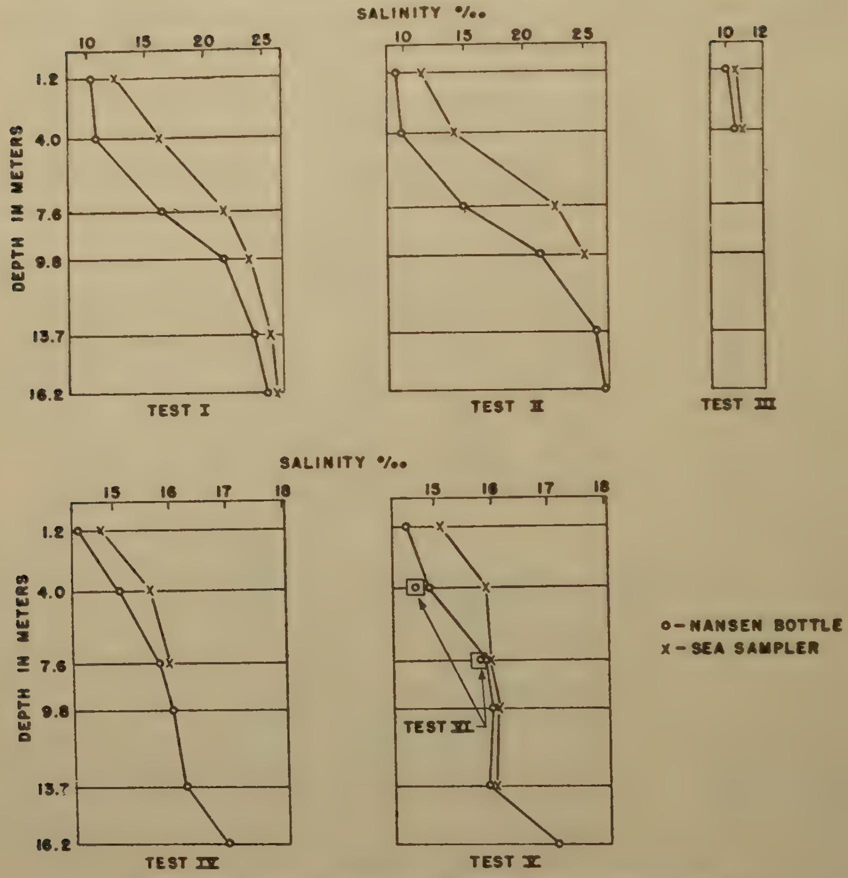


Figure 9. Tests of Sea Sampler Compared with Nansen Bottles

messenger was released to reverse the Nansen bottles at the same time as the sea sampler was withdrawn from the bottom.

In both these tests the sea sampler was brought up at the rate of about 60 feet in 10 seconds. Marked differences between the salinities were observed. These are shown in the curves in Fig. 9. The differences were systematic, however, and in the right sense to indicate that error was due to the throttling lag of the sampler bottles. Therefore, in test number 3 (Fig. 9), the sea sampler was brought up very slowly, with frequent stops to permit the complete interchange of water. It is immediately seen in this test that the differences between the Nansen bottles and the sea sampler are very materially reduced.

A further series, 4, 5, and 6 (Table IB), were made in about 70 feet of water just off Sandy Hook. Here the gradient of salinity was much less, varying from about 14½ °/₀₀ to 17 °/₀₀. Tests 4 and 5 were made with exactly the same procedure as tests 1 and 2. They show the same systematic excess of salinity obtained by the sea sampler over that of the corresponding Nansen bottle (Fig. 9). In this case, the difference is less because the gradient of salinity is less.

Immediately following test 5, the ship was put underway. As the test area was passed over at the speed of 8½ knots, the sea sampler was lowered for test 6. Here, the salinities obtained by the sea sampler were far closer to those of the Nansen bottles. This great reduction of the throttling error when the instrument is used underway is to be expected from two factors. First, in using the instrument underway, far more wire is played out to reach a particular depth; thus, in this case, it took 35 seconds to raise the instrument vertically through 60 feet of water compared to 10 seconds for the same depth at anchor. Second, with the vessel moving forward at the rate of 8½ knots, the relative velocity of the water through the bottles is much greater (even though the vertical speed of the bottle through the water is less than in a lowering made while hove to). These two factors both act to reduce the throttling error. The increased relative velocity of the water causes a greater rate of exchange in the bottle with time. The decreased vertical velocity of the bottle through the gradient causes a decreased rate of change of concentration with time, both factors tending to reduce the error. These effects are discussed quantitatively below.

THE LAG OF A SAMPLER DUE TO THROTTLING

In the pressure-operated sea sampler, a sample is taken by closing a container at a certain point as the whole instrument is moving through an environment with varying concentration of the constituent to be measured. In such a case there will be a certain error in the concentration determined from the entrapped sample due to imperfect exchange of the fluid through the container.

The diagram (Fig. 10) shows the arrangement of the cocks and the arrow shows the motion of the water relative to the sampler at a speed, v_r; then, if p_1 and p_2 are the pressures in front and at the back of the bottle respectively, the volume of water passing through the bottle in unit time is

$$Q = C_q \, a \sqrt{\frac{2(p_1 - p_2)}{\rho}},$$

where C_q is an orifice coefficient for the bottle, a is the area of the stop-

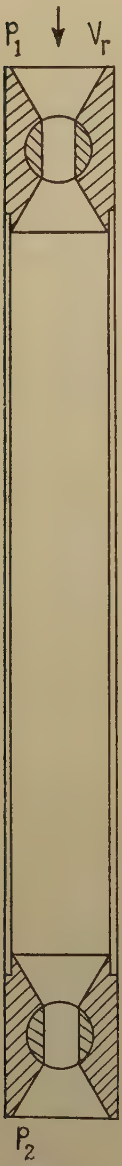


Figure 10. Diagram of Sampler Bottle

cock openings and ρ is the density of water. $p_1 - p_2$ will be proportional to the impact pressure of the water $= K\rho v_r^2/2$, where K is a constant less than unity due to any streamlining effect. Substituting,

$$Q = C_q \sqrt{K}\, a v_r \, .$$

Putting $C_q\sqrt{K} = k$, a constant for the instrument, the rate of mass flow into the container is

$$\frac{dM}{dt} = \rho k a v_r \, .$$

Let the concentration of the constituent be $C = M_s/M$, where M_s is the mass of the constituent and M is the mass of water. In an interval of time dt, suppose the sampler moves up a distance $-dz$, where the concentration in the environment has changed from C to $C - dC$. Then, if the concentration in the container was originally C_i, and complete mixing inside is assumed, an amount of water dM with this concentration has flowed out, and has been replaced by a similar amount dM with concentration

$$\frac{C + (C - dC)}{2} = C - \frac{dC}{2} \, .$$

The concentration in the container, assuming complete mixing[3], is then

$$C_i + dC_i = \frac{(M - dM)C_i + (C - dC/2)\, dM}{M}$$

$$dC_i = -\frac{(C_i - C)}{M}\, dM \, ,$$

so that the rate of change of concentration inside is then

$$\frac{dC_i}{dt} = -\frac{(C_i - C)}{M}\frac{dM}{dt} = -(C_i - C)\frac{V_r}{\xi} \, ,$$

---

[3] If *no* mixing is assumed, the rate of change of concentration inside is doubled, and hence all errors due to throttling are halved.

where $M = \rho V$ and $ka/V = 1/\xi$, which is a characteristic of the sampler bottle on the instrument.

Suppose, first, that the sampler is being raised through water which has a concentration C_o, constant for a considerable depth (so that the water in the container may be assumed also to be C_o). Further, that upon reaching a certain point, there is a sudden discontinuity of concentration which drops to C and remains constant with further decrease of depth. Integrating the equation for the rate of change of concentration in the container, we obtain

$$\frac{C_i - C}{C_o - C} = e^{-v_r t/\xi} \quad ,$$

if t is counted zero at the discontinuity. ξ/v_r is the lag coefficient of the sampler in units of time, similar to the lag coefficient of a thermometer (see for example Middleton, 1941: 56–58). ξ/v_r represents the time for the difference of concentration in the container C_i, and in the environment C, to reduce itself to $1/\epsilon$ of its original value.

When the concentration varies with depth z, measured positive downwards with a gradient $\alpha = dc/dz$, then, if the whole sampler is moving upward at a rate $v_v = -\,dz/dt$, the concentration changes at a rate

$$\frac{dc}{dt} = -\,v_v\alpha\,.$$

Then

$$C = C_o - V_v\alpha t,$$

where C_o is the concentration at some lower point from which t is counted zero. The rate of change of concentration in the container is then

$$\frac{dC_i}{(C_i - C_o + v_v\alpha t)} = -\,\frac{v_r}{\xi}\,dt\,,$$

or

$$\frac{dC_i}{dt} = -\,\frac{V_r}{\xi}\left[C_i - (C_o - v_v\alpha t)\right],$$

and

$$C_i - C = -\,\xi\alpha\left(\frac{v_v}{v_r}\right)\left(1 - e^{-v_r t/\xi}\right),$$

after $t >> \xi/v_r$

$$C_i - C = -\,\xi\alpha\left(\frac{v_v}{v_r}\right) = -\,\frac{V}{ka}\,\alpha\left(\frac{v_v}{v_r}\right).$$

If the sampler is being used from a vessel at anchor in no current, then $-v_v = v_r$,

and
$$\frac{C_i - C}{\alpha} = \frac{V}{ak} = \xi \, .$$

If, however, it is being towed at a forward speed, then $-(v_v/v_r) = \cos\beta$, where β is the inclination of the path of the instrument to the vertical, or approximately the wire angle,[4] then

$$\frac{C_i - C}{\alpha} = \xi \cos\beta \, .$$

In a constant gradient of concentration, the indicated gradient, α_i, will equal the actual gradient α. Thus, the expression shows that if the characteristic ξ of the sampler is known, and if β is known, the indications may be corrected if necessary.

THE LAG COEFFICIENT AND "CHARACTERISTIC LENGTH" OF A SAMPLER

As the lag coefficient ξ/v_r of a sampler depends on the relative speed of the instrument to the water, it is more convenient to discuss lag in terms of ξ, which is a characteristic of the instrument alone. ξ has the dimensions of length. An ideal sample bottle with no obstruction to flow through it would have k = unity. Then $\xi = V/a$, and as the area a would be the area of cross section of the bottle for no obstruction, ξ would represent the actual length of this ideal bottle. In the case of real bottles with stop-cocks, k must always be less than unity, and it is then convenient to define ξ as the "equivalent length" of an ideal bottle. Equivalent length then is the length of a bottle with no throttling, which would have the same lag. If we assume that the tripping depth is at the depth of the top of the bottle, then in a steady gradient of salinity the actual sample in the bottle will always represent the salinity at the bottom of the equivalent ideal bottle. Thus

$$C_i = C + \alpha\xi\cos\beta \text{ (see Fig. 11).}$$

---

[4] β is only equal to the wire angle δ with the vertical, when the drag of the wire is considered negligible and when the speed of hauling in wire, v_w, is negligible compared to the forward speed of the vessel, v_f. In other cases, β is always larger than δ. If the assumption is made that the sea sampler lines up with the direction of the flow of water relative to it, then if the instrument has a drag equal to b times the square of the relative velocity v_r, and a weight W, it can be shown that $\cos\beta = b/w \, v_w v_f \cos\delta/\tan\delta$.

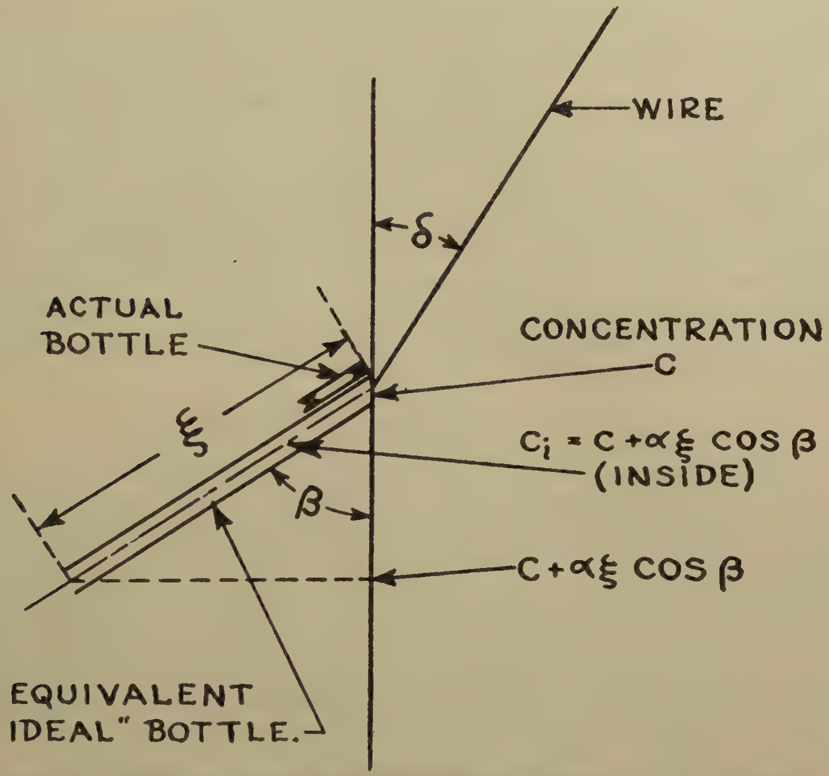

WIRE
δ
ACTUAL BOTTLE
CONCENTRATION
C
ξ
$C_i = C + \alpha\xi \cos\beta$
(INSIDE)
β
$C + \alpha\xi \cos\beta$
EQUIVALENT "IDEAL" BOTTLE.

Figure 11. Equivalent Length of the Sampler Bottle in Relation to Lag in a Constant Gradient of Concentration

EXPERIMENTAL DETERMINATION OF THE EQUIVALENT LENGTH OF THE SEA SAMPLER

From tests 1, 2, 4 and 5, which were performed with the vessel stationary, $(C_i - C)/\alpha$ has an average value of about 2.5 m. (Fig. 9).[5] Therefore, 2.5 m. is the equivalent length of the present sea sampler. When used underway at 8½ knots, however, the inclination of the bottle to the vertical is very great and the actual vertical distance between the top and bottom of the equivalent ideal bottle is reduced by the factor of the cosine of this inclination, so that the error due to throttling is reduced in this case to a very small fraction of the 2.5 m. which obtains for stationary operation.

---

[5] This value was computed in the following manner:

In tests 1 and 2 (Fig. 9) the gradient α was computed between the depths 9.8 m. and 4 m. where it is fairly constant. The error $(C_i - C)$ was measured at 4 m., assuming that there had been time enough to attain constant lag (5.8 m. $> \xi$). In the case of tests 4 and 5 (Fig. 9), the gradient between 7.6 m. and 1.2 m. and the error, at 1.2 m., were utilized. The value of the equivalent length from tests 4 and 5 agrees fairly well with that for tests 1 and 2, which were in different water.

Apart from the factor K in the expression $\xi = V/aC_q\sqrt{K}$, all the factors are easily determinable in the laboratory. C_q, determined by a discharge test at various heads of water through the sampler bottle totally immersed, was found to be about 0.57, $V = 120$ cm^2 and $a = 1.44$ cm^2. Using $\xi = 2.5 + 10^2$ cm gives $K = 0.34$. This means that the streamlining of the whole instrument is such that only about one-third of the full dynamic pressure on the bottle is effective in causing the interchange of water.

PRACTICAL USE OF SEA SAMPLER

An example of the use of the sea sampler was provided when Woods Hole Oceanographic Institution was requested to investigate a problem concerning the dispersal of acid waste as it was released from a barge. The BALANUS followed in the wake of the barge at various distances, and about 50 lowerings with the sampler were made; a vertical section was thereby constructed showing the concentration of the waste at various horizontal distances and depths from the point of efflux in the barge. It is cited as an example of an unusual type of operation that can be accomplished with the sea sampler which would hardly be possible by use of other existing means.

SUMMARY

An improved model of the thermographic sea sampler (a sea sampler combined with a conventional bathythermograph) is described. It consists of twelve sample bottles attached to a standard bathythermograph and arranged to trip at various pressures (depths) which can be chosen at will. Comparisons of samples taken in this manner with those obtained by the use of Nansen bottles are shown and a theory for the lag of a sampling device of this kind is given. Theory and observation indicate that the errors of the sea sampler are small when used underway. The theory further permits correction of observations for lag when necessary in terms of a single "characteristic length" of the instrument and the angle of inclination of its path through the water.

REFERENCES

MIDDLETON, W. E. K.

1941. Meteorological instruments. Toronto Univ. Press. 213 pp.

SPILHAUS, A. F.

1938. A bathythermograph. J. Mar. Res., *1* (2): 95–100.

1940. A detailed study of the surface layers of the ocean in the neighborhood of the Gulf Stream with the aid of rapid measuring hydrographic instruments. J. Mar. Res., *3* (1): 51–75.

1941. Fine structure of the edge of the Gulf Stream. Trans. Amer. geophys. Un., *22:* 478–484.

THE THEORY OF THE ELECTRIC FIELD INDUCED IN DEEP OCEAN CURRENTS[1]

By

HENRY STOMMEL

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts

The Problem. Due to the fact that the water in the ocean is a conductor, and that it is everywhere under the influence of the earth's magnetic field, we should expect, by the law of electric induction, that wherever the water is in motion electric potentials and currents will be established. It is the purpose of this paper to discuss the theory of these phenomena in the deep ocean and to indicate certain analytical solutions of the problem which demonstrate the important physical aspects of electric fields associated with ocean currents.

The Fundamental Equation of the Electric Field.[2] Let $\mathbf{H}$ be the vector magnetic field intensity, ρ the scalar resistivity and $\mathbf{i}$ the electric current vector. Consider a closed curve fixed in the fluid. Let $\mathbf{v}$ be the fluid velocity vector. The length of arc along the closed curve is given by the vector $\mathbf{s}$. The element of area of the surface enclosed by the curve is $\mathbf{dS}$. Faraday's law of induction may then be expressed in the following general form,

$$\frac{d}{dt}\int\int \mathbf{H} \cdot \mathbf{dS} + \oint \mathbf{i} \cdot \mathbf{ds} = 0 \ , \qquad (1)$$

where the d/dt is to be understood in the manner of the substantial derivative of hydrodynamics. The first equation (1) may be transformed (Abraham and Becker, 1932: 39–40) as follows:

$$\frac{d}{dt}\int\int \mathbf{H} \cdot \mathbf{dS} = \int\int \frac{\partial \mathbf{H}}{\partial t} \cdot \mathbf{dS}$$
$$+ \int\int (\text{div}\ \mathbf{H})\mathbf{v} \cdot \mathbf{dS} - \int\int \Big[\text{curl}\ (\mathbf{v} \times \mathbf{H})\Big] \cdot \mathbf{dS} \ . \qquad (2)$$

---

[1] Contribution No. 453 from the Woods Hole Oceanographic Institution.

[2] The author is particularly indebted to Mr. M. S. Longuet-Higgins of the Admiralty Research Laboratory at Teddington, England, for acquainting him with the basic physics involved in this section; see REFERENCES.

In the ocean a considerable simplification of (2) is possible, because over moderately large areas **H** is constant and uniform. Therefore, the first two terms of the second member vanish, and by Stoke's Theorem the third term may be transformed to a line integral,

$$-\iint \left[\operatorname{curl} (\mathbf{v} \times \mathbf{H}) \right] \cdot d\mathbf{S} = -\oint (\mathbf{v} \times \mathbf{H}) \cdot d\mathbf{s}. \tag{3}$$

Equation (1) may now be written in the following form:

$$\oint (\mathbf{v} \times \mathbf{H} - \rho\mathbf{i}) \cdot d\mathbf{s} = 0. \tag{4}$$

The vanishing of this line integral signifies the existence of the electric scalar potential function ϕ, defined in the following way:

$$\nabla\phi = \mathbf{v} \times \mathbf{H} - \rho\mathbf{i}. \tag{5}$$

The term $\rho\mathbf{i}$ involving the unknown current vector **i** may be eliminated if ρ is assumed to be uniform, for then the div $\rho\mathbf{i}$ vanishes, and upon taking the divergence of (5) one obtains

$$\boxed{\nabla^2\phi = \mathbf{H} \cdot \operatorname{curl} \mathbf{v}.} \tag{6}$$

If **H** and **v** are both regarded as known, which is physically the case, then we simply have here Poisson's equation to solve for the electric potential ϕ. It is desirable to emphasize the complete generality of (6). This equation defines the electric field resulting from *any arbitrary velocity field* in the ocean.

The boundary conditions upon this equation are obtained from (5). At a nonconducting boundary the normal component of the vector **i** vanishes. If **n** is the unit vector normal to the boundary,

$$\mathbf{n} \cdot \mathbf{i} = 0,$$

so that

$$\frac{\partial\phi}{\partial\mathbf{n}} = \mathbf{n} \cdot \mathbf{v} \times \mathbf{H}.$$

At a conducting boundary, such as occurs between two media of different specific resistivity, the normal component of **i** is equal in each medium, and the potential ϕ is equal in each medium. The first condition may be expressed in terms of $\partial\phi/\partial\mathbf{n}$ for each medium by using (5).

An Idealized Particular Ocean Current System and its Associated Electric Field. The velocity and potential fields of a natural ocean current system are very complicated functions of position. They may be treated numerically as indicated in the section, Numerical Method of Solution. An analytically simple model is advantageous, inasmuch as it shows the essential features without many complications of a purely formal but not conceptual nature. In order to exhibit some of the features of the problem, an idealized ocean current system is treated here. Rectangular co-ordinates are taken with the x,y-plane in the surface of the ocean, with the z-axis pointing vertically upwards. This ocean is supposed to be divided into two layers. The top layer, no. 1, extends from $z = 0$ to $z = -h_1$. The bottom layer, no. 2, extends from $z = -h_1$ to $z = -h_2$. The bottom of the ocean, $z = -h_2$, is taken in this discussion as nonconducting.

The velocity in layer no. 1 is supposed to be given as

$$\left\{\begin{matrix} v_x \\ v_y \\ v_z \end{matrix}\right\} = \left\{\begin{matrix} 0 \\ v_o \cos \pi/b\, x \\ 0 \end{matrix}\right\}. \tag{7}$$

Layer no. 2 is supposed to be at rest. This current system consists of alternate bands of water of breadth b moving in the positive and negative y-direction over a lower layer which does not participate in the motion. This picture is artificial, but may be generalized by implementing the Fourier Integral Theorem.

Let us suppose that the magnetic field of the earth is entirely vertical,

$$\mathbf{H} = \left\{\begin{matrix} 0 \\ 0 \\ H_z \end{matrix}\right\}. \tag{8}$$

Substitution of (7) into (6) gives the equations

$$\nabla^2\phi_1 = \gamma \sin \beta\, x, \tag{9}$$

$$\nabla^2\phi_2 = 0, \tag{9'}$$

where the subscript refers to the layer number, and the following abbreviated forms have been introduced:

$$\gamma = -H_z v_o\, \pi/b\,, \qquad \beta = \pi/b\,. \tag{10}$$

In order to obtain a solution of (9) and (9′), one supposes that the potential is expressible in separable form, which indeed later appears to be the case.

$$\phi_1 = \sin \beta x \cdot Z_1\,, \tag{11}$$

$$\phi_2 = \sin \beta x \cdot Z_2\,. \tag{11'}$$

Substitution of (11) and (11′) into (9) and (9′) yields, after cancelling the common factor,

$$\frac{\partial^2 Z_1}{\partial z^2} - \beta^2 Z_1 = \gamma\,, \tag{12}$$

$$\frac{\partial Z_2}{\partial z^2} - \beta^2 Z_2 = 0\,. \tag{12′}$$

The solutions of these two equations are simply,

$$Z_1 = A_1 \cosh \beta z + B_1 \sinh \beta z - \delta\,, \tag{13}$$

$$Z_2 = A_2 \cosh \beta z + B_2 \sinh \beta z\,, \tag{13′}$$

where A_1, A_2, B_1, and B_2 are constants of integration and

$$\delta = \beta^{-2}\gamma\,.$$

These constants must now be determined in terms of the boundary conditions. Take **n** as the unit vertical vector. At $z = 0$ there must be no current across the water surface:

$$\mathbf{i} \cdot \mathbf{n} = 0\,. \tag{14}$$

From (5) we have the electric current in terms of the potential

$$\mathbf{i} = \frac{1}{\rho}\left[\mathbf{v} \times \mathbf{H} - \nabla\phi\right]. \tag{15}$$

Substituting into (14) we obtain

$$\mathbf{n} \cdot \mathbf{v} \times \mathbf{H} - \frac{\partial \phi}{\partial z} = 0\,. \tag{16}$$

In the particular case we are discussing $\mathbf{n} \cdot \mathbf{v} \times \mathbf{H}$ vanishes, so that the boundary condition at the surface $z = 0$ is

$$\frac{\partial \phi}{\partial z} = 0\,. \tag{17}$$

The boundary condition at the bottom $z = -h_2$ is also one of no current, so it is the same as (17).

At the interface between layers 1 and 2, $z = -h_1$, two conditions prevail, for clearly the potential and current must be continuous. Therefore at $z = -h$,

$$\phi_1 = \phi_2\,, \tag{18}$$

$$\mathbf{i}_1 = \mathbf{i}_2\,. \tag{18′}$$

The second of these equations may be expressed in terms of the potential function

$$\frac{\partial\phi_1}{\partial z} = \frac{\partial\phi_2}{\partial z}. \tag{18''}$$

The conditions (17) at $z = 0$ and $z = -h_2$, and the conditions (18) and (18″) at $z = -h_1$, suffice to determine the constants of integration.

$$\begin{aligned} A_1 &= \delta\frac{\sinh\beta(h_2 - h_1)}{\sinh\beta h_2}, \\ B_1 &= 0 \\ A_2 &= -\delta\frac{\sinh\beta h_1}{\tanh\beta h_2}, \\ B_2 &= -\delta\sinh\beta h_1. \end{aligned} \tag{19}$$

Therefore, the final solution is

$$\phi_1 = \sin\beta x\left[1 - \frac{\sinh\beta(h_2 - h_1)}{\sinh\beta h_2}\cosh\beta z\right]\frac{v_oH_z}{\beta}, \tag{20}$$

$$\phi_2 = \sin\beta x\left[\frac{\sinh\beta h_1}{\tanh\beta h_2}\cosh\beta z + \sinh\beta h_1\sinh\beta z\right]\frac{v_oH_z}{\beta}. \tag{20'}$$

The exact expressions (20) and (20′) may be replaced by simpler expressions to a first approximation in the case where $\beta h_2 << 1$, that is, where the breadth of the current is much greater than the depth of the ocean, as most frequently happens in nature.

$$\phi_1 \cong \phi_2 \cong \sin\beta x\left[\frac{h_1}{h_2}\right]\frac{v_oH_z}{\beta}. \tag{21}$$

According to this approximate formula the lines of equipotential are vertical straight lines, and the amplitude of the potential developed is directly proportional to the ratio of current depth to ocean depth.

For shallow surface currents the return circuit beneath, in the resting water, is comparatively so large that it effectively shorts the potentials developed.

Numerical Method of Solution. A convenient numerical method of solution may be based upon the relaxation methods developed by Southwell (1946). Any vertical velocity profile can be investigated and the electric potential field computed.

Equation (6) is expressed in finite difference form. The boundary condition at the bottom and at the free surface is expressed by (16). The procedure of the solution is then by liquidation of residuals.

Publication of any of these relaxation solutions is postponed until some *simultaneous* electric potential measurements and sufficient hydrographic stations for computation of the velocity field are available.

The Effect of the Resistance of the Bottom. Longuet-Higgins and Barber (1946) have shown that the potential developed in a tidal current through a shallow channel (such as the English Channel) is greatly controlled by the resistivity of the bottom. It is important, therefore, to discover what the effect of the electric conduction through the bottom will be on the preceding analysis, which was made upon the assumption of a nonconducting bottom.

For this purpose we may introduce a layer 3 extending from $z = -h_2$ to $z = -\infty$ which will represent the bottom. The resistivity of layers 1 and 2 is ρ_1 and ρ_2, and since they are bottom sea water layers $\rho_1 = \rho_2$. The resistivity ρ_3 of the bottom is considerably greater than ρ_2. For convenience we introduce the ratio

$$\tilde{\omega} = \frac{\rho_2}{\rho_3}; \qquad 1 >> \tilde{\omega} \geq 0 .$$

The boundary conditions at $z = 0$ and $z = -h_1$ remain the same as before. For $z = -h_2$ however we now have

$$\phi_2 = \phi_3 , \tag{22}$$

$$\frac{\partial \phi_2}{\partial z} = \tilde{\omega} \frac{\partial \phi_3}{\partial z} . \tag{22$'$}$$

At $z = -\infty$ the potential vanishes

$$\phi_3 = 0 \text{ at } z = -\infty . \tag{23}$$

The governing equations (13) and (13′) still hold, with the addition of

$$Z_3 = A_3 \cosh \beta z + B_3 \sinh \beta z . \tag{23$''$}$$

The condition (23) immediately determines B_3, so that we have

$$Z_3 = A_3 (\cosh \beta z - \sinh \beta z) .$$

Again, at $z = 0$ the condition (17) shows that $B_1 = 0$.

There remain four independent constants A_1, A_2, B_2, A_3, to determine and the four boundary conditions (18), (18″), (22) and (22′). These give respectively

$$\begin{aligned} A_1 \cosh \beta h_1 - A_2 \cosh \beta h_1 + B_2 \sinh \beta h_1 &= \delta , \quad (24) \\ - A_1 \sinh \beta h_1 + A_2 \sinh \beta h_1 - B_2 \cosh \beta h_1 &= 0 , \\ A_2 \cosh \beta h_2 - B_2 \sinh \beta h_2 - A_3 (\cosh \beta h_2 + \sinh \beta h_2) &= 0 , \\ - A_2 \sinh \beta h_2 + B_2 \cosh \beta h_2 + A_3 \tilde{\omega} (\sinh \beta h_2 + \cosh \beta h_2) &= 0 . \end{aligned}$$

For the purpose of determining these constants it will be convenient to restrict our thoughts to the case where $\beta h_2 << 1$, that is, broad currents. The above equations are then, to a first approximation, as follows:

$$\begin{aligned} A_1 - A_2 + B_2\,\beta h_1 + 0 &= \delta\,, \\ -A_1\,\beta h_1 + A_2\,\beta h_1 - B_2 + 0 &= 0\,, \\ 0 + A_2 - B_2\,\beta h_2 - A_3\,(1 + \beta h_2) &= 0\,, \\ 0 - A_2\,\beta h_2 + B_2 + A_3\,\tilde{\omega}\,(1 + \beta h_2) &= 0\,. \end{aligned} \tag{25}$$

The values of the constants are determined to be (neglecting second and higher powers of βh_1 and βh_2)

$$\begin{aligned} A_1 &= \delta\,\frac{\tilde{\omega} + \beta\,(h_1 - h_2)}{\tilde{\omega} - \beta\,h_2} = \frac{\beta\,(h_2 - h_1) - \tilde{\omega}}{\beta\,h_2 - \tilde{\omega}} \\ A_2 &= -\,\delta\,\frac{\beta h_1}{\beta h_2 - \tilde{\omega}}\,, \\ B_2 &= -\,\delta\beta h_1, \\ A_3 &= \frac{\delta\,\beta h_1}{(1 + \beta h_2)\,(\tilde{\omega} - \beta h_2)}\,. \end{aligned} \tag{26}$$

Comparison of these values of A_1, A_2 and B_2 with those obtained for the nonconducting bottom ($\tilde{\omega} = 0$) shows that for small $\tilde{\omega}$ ($\tilde{\omega} << \frac{1}{10}$) the potential developed is affected by the presence of the bottom only negligibly. These considerations make it seem justifiable to us to ignore the resistivity of the bottom in discussion of the potential fields generated by shallow currents in the deep ocean. In the case of currents in shallow channels these conclusions do not apply.

Summary. The physical principles which describe the association of the water velocity and induced electric potential field in the deep ocean are formulated, and with the aid of a simple idealized model the chief features are discussed.

REFERENCES

Abraham, Max and Richard Becker
1932. Classical electricity and magnetism. English translation. Blackie and Son, London. 285 pp.
Longuet-Higgins, M. S., and N. F. Barber
1946. The measurement of water velocity by electromagnetic induction. Admiralty Res. Lab. Rep. R. 1/102.22/W.
Southwell, R. V.
1946. Relaxation methods in theoretical physics. Oxford Univ. Press, London. 248 pp.

SOUND SCATTERERS IN THE SEA[1]

By

RUSSELL W. RAITT

Marine Physical Laboratory
University of California
San Diego 52, California

INTRODUCTION

When sound pulses of ultrasonic frequency are transmitted into the sea, many echoes are returned to the source from scatterers distributed throughout the sea. Under the usual conditions of observation with echo-sounding and echo-ranging apparatus, the echoes merge indistinguishably. The resulting irregularly decaying sound is called reverberation (Eyring, *et al.*, 1948) by analogy with the reverberation observed in rooms.

Owing to the failure to observe discrete echoes from single scatterers in the usual reverberation records, studies of the properties of reverberation are customarily confined to measurement of the aggregate scattering power of large numbers of scatterers. These studies have yielded useful information (see References *1–8*) regarding the distribution in space and time of the average scattering power of the sea, but the problem of identification of the scatterers remains unsolved. In considerable degree this failure is due to the lack of knowledge of the size of the individual scatterers.

ECHOES FROM SINGLE SCATTERERS

The reverberation records do not always exhibit the complex character ascribable to many overlapping echoes. If the pulse length is made very short and if the concentration of scatterers is low, it is possible to observe single echoes. Fig. 1 represents an example of such a case. It is an oscillogram taken on the USS Jasper during a study of reverberation by the Echo Ranging Section of the University of California, Division of War Research, in March and April 1945. The sound projector was suspended overside at a depth of 5 m and directed vertically downward. The technique of this type of observation is described elsewhere (Eyring, 1948).

---

[1] This work represents one of the results of research carried out under contract with the Bureau of Ships, Navy Department.

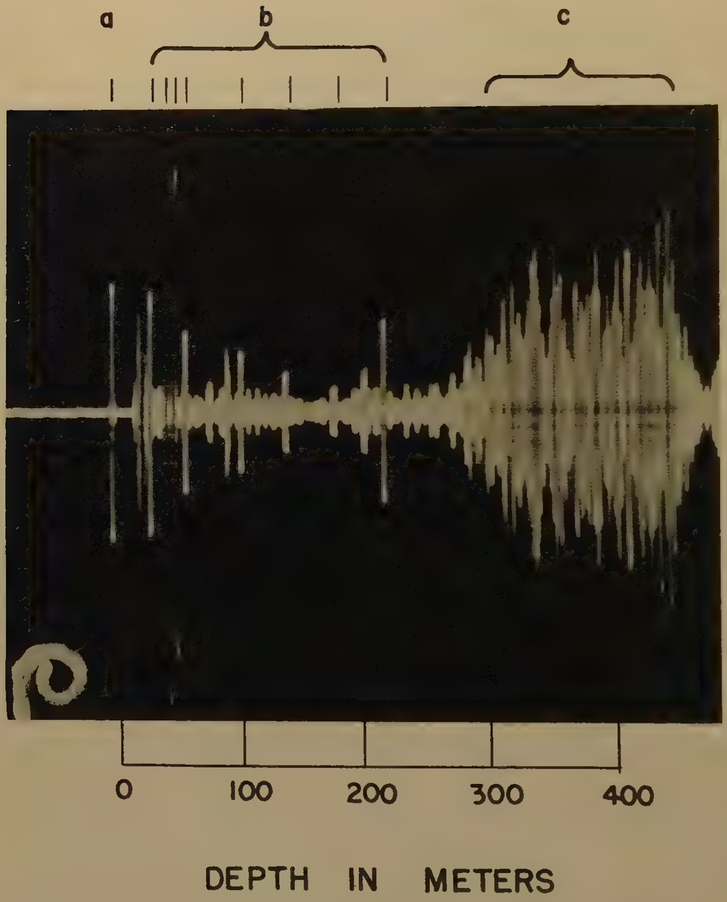


Figure 1. Oscillogram illustrating echoes from scatterers in the sea. Indicated events are: (a) transmitted pulse of 24 kc, 3 ms duration; (b) echoes from scatterers; (c) reverberation from a layer of strong concentration of scatterers.

Significant events on the oscillogram are identified by letters: (a) transmitted pulse of 24 kc frequency and 3 ms duration; (b) echoes from single scatterers; (c) reverberation from a strongly scattering layer between 300 and 450 m depth. It is to be noted that the echoes (b) have the duration of the transmitted pulse and have a very sharp beginning and ending, in marked contrast to the ragged and irregular "spikes" of the typical reverberation. Furthermore, in the region of the echoes, the record differs in character from the region of the scattering layer. Many small echoes, incompletely resolved, can be seen.

The illustrated oscillogram was recorded on a cruise conducted off the coast of Lower California from San Diego to the Gulf of Lower California. At 13 stations distributed over this region, sets of 10 to 20 oscillograms were recorded at a frequency of 24 kc and pulse length of 3 ms. On most of these sets, echoes similar to those illustrated in Fig. 1 were observed. All of the records were systematically examined and the echo levels were recorded. Before presenting the results of this analysis it is desirable to discuss the problem of the observation of echoes from a distribution of scatterers.

THEORY OF ECHOES FROM A DISTRIBUTION OF SCATTERERS

It is assumed that the medium has a constant sound velocity and no absorption. Later, effects of transmission loss can be introduced. Let e_1 be the voltage amplitude of the transmitted pulse received at the hydrophone terminals when the hydrophone is faced towards the projector at a distance of one meter. Actually e_1 is a fictitious quantity usually determined in an indirect manner with the assistance of auxiliary transducers or with a standard target. Let e be the voltage amplitude at the hydrophone terminals produced by an echo. If a single scatterer is situated on the sound beam axis at a distance of R meters, its scattering cross section σ, in square meters, is defined by the equation

$$(e/e_1)^2 = \sigma/4\pi R^4. \qquad (1)$$

Physically, the scattering cross section σ is the effective area of the sound beam intercepted by the sound scatterer. It is not necessarily equal to the actual cross section. If a scatterer is semitransparent to sound, or if it is opaque but much smaller than the wavelength of the incident sound, the acoustic cross section will be smaller than the actual cross section. In the special case of a rigid sphere of radius greater than the wavelength the acoustic cross section equals the area intercepted by the sphere.

It is convenient in analysis of the data to use the quantity $E = (4\pi R^4 e^2/e_1^2)$ as a measure of echo intensity. For a scatterer on the sound beam axis, $E = \sigma$. If the scatterer is not on the axis,

$$E = \sigma q^2\,(\theta,\ \phi), \tag{2}$$

where $q\ (\theta,\ \phi)$ is the transducer directivity pattern expressed as a function of the polar angle θ and aximuthal angle ϕ. If the transducer is circularly symmetrical, q is a function of θ only.

Uniform Distribution of Scatterers of Equal Cross Section

In general, the sound scatterers in the sea cover a wide range of variation in size and shape and are irregularly distributed. Even if it were possible to determine the position and cross section of each individual scatterer within the field of observation, the complexity of the analysis would be formidable. However, from the oscillograms, only the range and echo intensity of a scatterer are known; the position relative to the beam axis is unknown and hence its cross section is indeterminate. The problem is necessarily statistical.

As a first step, let it be assumed that the sound beam is directed into a medium uniformly filled with scatterers, all having the same cross section σ. Let ν be the number per unit volume. On the average the number of echoes received from distances between R and $R + dR$ is equal to $4\pi\nu R^2 dR$. These echoes will be distributed in intensity from a maximum value of $E = \sigma$ for scatterers on the sound beam axis to negligibly small values for scatterers well outside the beam. Owing to the fact that the solid angle included within the strong central portion of the sound beam is much less than the total solid angle insonified, the weak echoes will be much more numerous than the strong ones.

For convenience, let $\eta = E/\sigma$ and let $f\ (\eta)\ d\eta$ be the fraction of echoes whose η lies between η and $\eta + d\eta$. The form of $f\ (\eta)$ is determined numerically from the measured directivity function $q^2\ (\theta,\ \phi)$. The value of $f\ (\eta)\ d\eta$ is equal to the fraction of the total solid angle included between the contours $q^2\ (\theta,\ \phi) = \eta$ and $q^2\ (\theta,\ \phi) = \eta + d\eta$.

The form of $f\ (\eta)$ for the sound projector used in the reported studies is shown in Fig. 2. It shows a strong increase of $f\ (\eta)$ towards decreasing η and illustrates quantitatively the above-mentioned prevalence of weak echoes relative to the strong ones. The cumulative distribution $\int_{\eta}^{1} f\ (\eta)\ d\eta$ is shown in Fig. 3. The ordinates give the fraction of echoes whose η is greater than the values designated by the abscissae.

When $f\ (\eta)$ is known, the distribution of echo intensities is readily

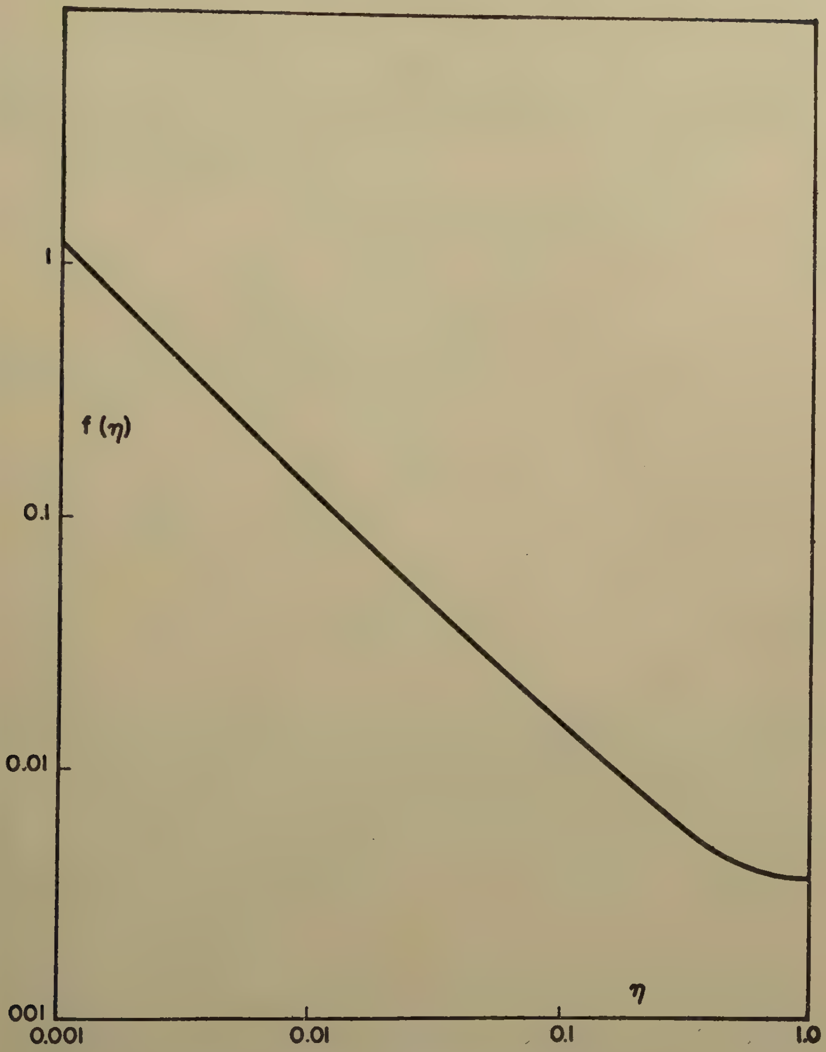


FIGURE 2. The directivity distribution function, $f(\eta)$.

obtained. Let $\phi(E, R)\,dE\,dR$ equal the number of echoes with echo intensities between E and $E + dE$ and distances between R and $R + dR$. Then $\phi(E, R)\,dE\,dR$ is the product of the total number of echoes received at distances between R and $R + dR$ and the fraction $f(\eta)\,d\eta$. Noting that $\eta = E/\sigma$,

$$\phi(E, R) = 4\pi\nu R^2 f(E/\sigma)/\sigma. \tag{3}$$

A useful form is the cumulative distribution, $\Phi(E, R)$, which gives the number of echoes with intensity greater than E.

$$\begin{aligned}\Phi(E, R) &= \int_E^\sigma \phi(E, R)\,dE \qquad (4)\\ &= 4\pi\nu R^2 \int_{E/\sigma}^1 f(E/\sigma)\,d(E/\sigma)\\ &= 4\pi\nu R^2 \int_{E/\sigma}^1 f(\eta)\,d\eta.\end{aligned}$$

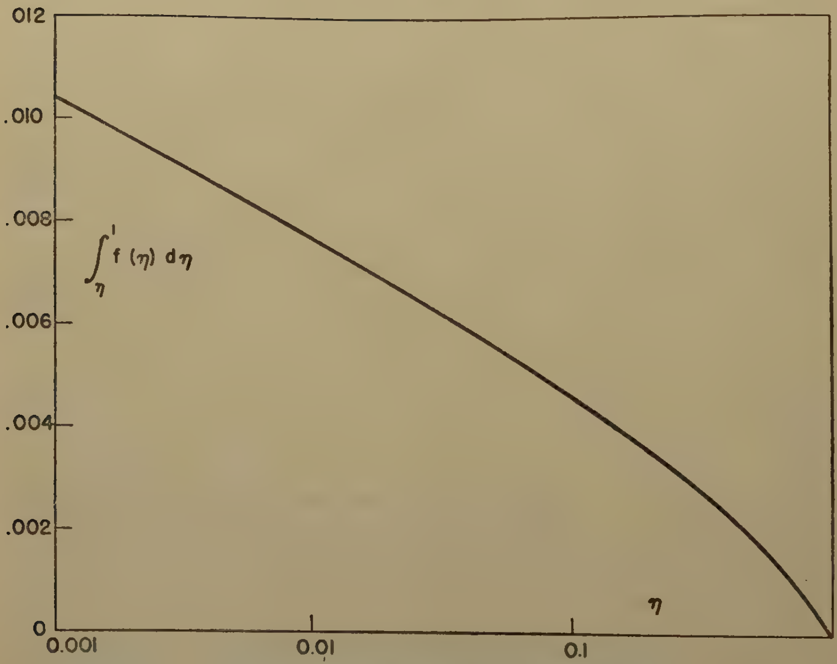


FIGURE 3. The cumulative directivity distribution function, $\int_\eta^1 f(\eta)\,d\eta$.

Scatterers Having a Distribution of Cross Sections

Let the scatterers have a distribution of cross sections specified by the function $\nu(\sigma)$, where $\nu(\sigma)\,d\sigma$ represents the number of scatterers per cubic meter with cross section between σ and $\sigma + d\sigma$. The echo distribution for this case is obtained by substituting $\nu(\sigma)\,d\sigma$ for ν in (3) and integrating over σ.

Thus the resulting distribution is

$$\phi(E, R) = 4\pi R^2 \int_E^\infty [\nu(\sigma) f(E/\sigma)/\sigma]\, d\sigma\,. \tag{5}$$

The cumulative distribution is obtained by substituting $\nu(\sigma)\,d\sigma$ for ν in (4), and integrating over σ. Let

$$F(\eta) = \int_\eta^1 f(\eta)\, d\eta\,.$$

Then,

$$\Phi(E, R) = 4\pi R^2 \int_E^\infty \nu(\sigma) F(E/\sigma)\, d\sigma\,. \tag{6}$$

In principle, if $\phi(E, R)$ were determined with precision, $\nu(\sigma)$ could be obtained by solution of (5). Unfortunately, only the stronger echoes are resolved on available oscillograms and hence our knowledge of the distribution of cross sections is necessarily incomplete. However, means are provided for estimating the size and concentration of the larger scatterers from the echo measurements. It is therefore of interest to examine the experimental results in the light of the above treatment.

OBSERVED DISTRIBUTION OF ECHO INTENSITIES

It was mentioned in a previous section that a large sample of oscillograms bearing numerous resolvable echoes was obtained during a cruise in March and April 1945. Each of these echoes was measured in a microfilm reader, and the echo intensities and depths of the scatterers were determined. Since the range of intensity values was large, they were converted to a *db* scale for convenience in analysis. On this scale the echo level, L, is given by

$$L = 10 \log E. \tag{7}$$

The echo levels of a set of 12 oscillograms, taken on March 12, 1945 at Lat. 28° 30′ N, Long. 118° 10′ W at 0914, Pacific Standard Time, are plotted as a function of depth in Fig. 4. The points shown as open circles represent clear cut, well resolved echoes; the solid circles represent echoes showing some evidence of interference by smaller

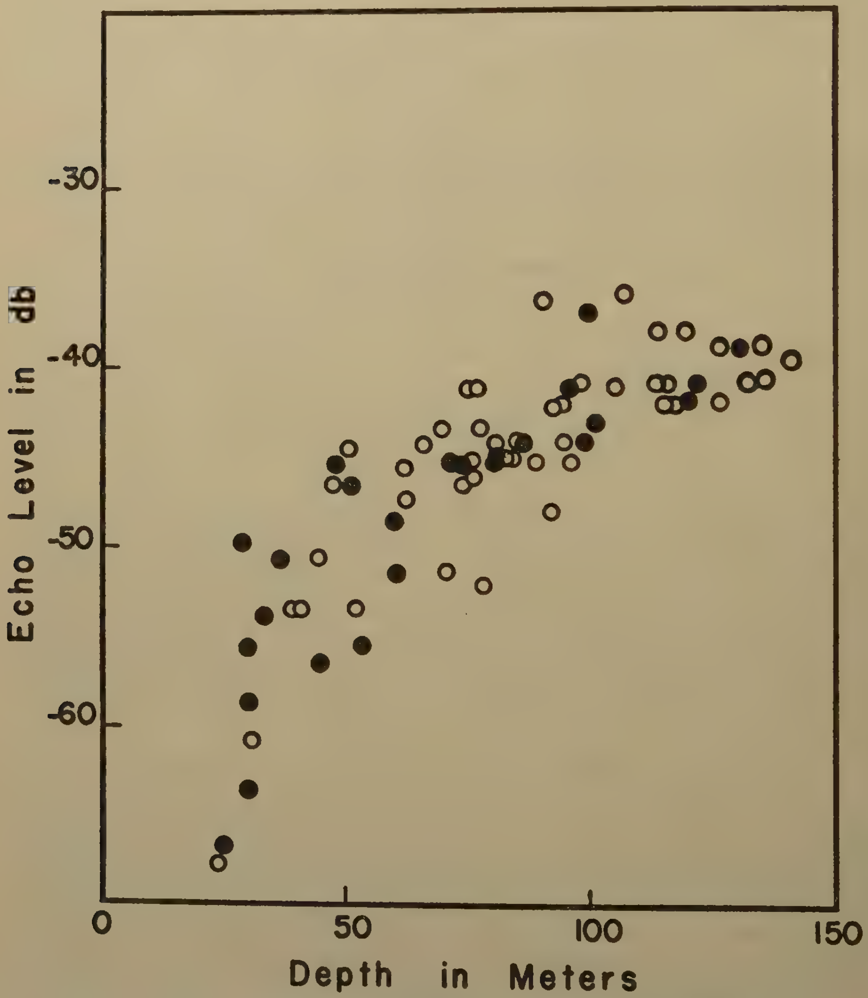


Figure 4. Echo levels recorded on March 12, 1945 at Lat. 28° 30′ N, Long. 118° 10′ W, at 0914 Pacific Standard Time.

ones, yet still giving good evidence that a single echo was the dominant event.

In interpreting these results it must be borne in mind that only the stronger echoes have been measured. Hence, only the region of the plot near the upper bound of the area occupied by points is significant. The lower bound is limited by the general level of unresolved reverberation. The increase of the upper bound with depth shows that

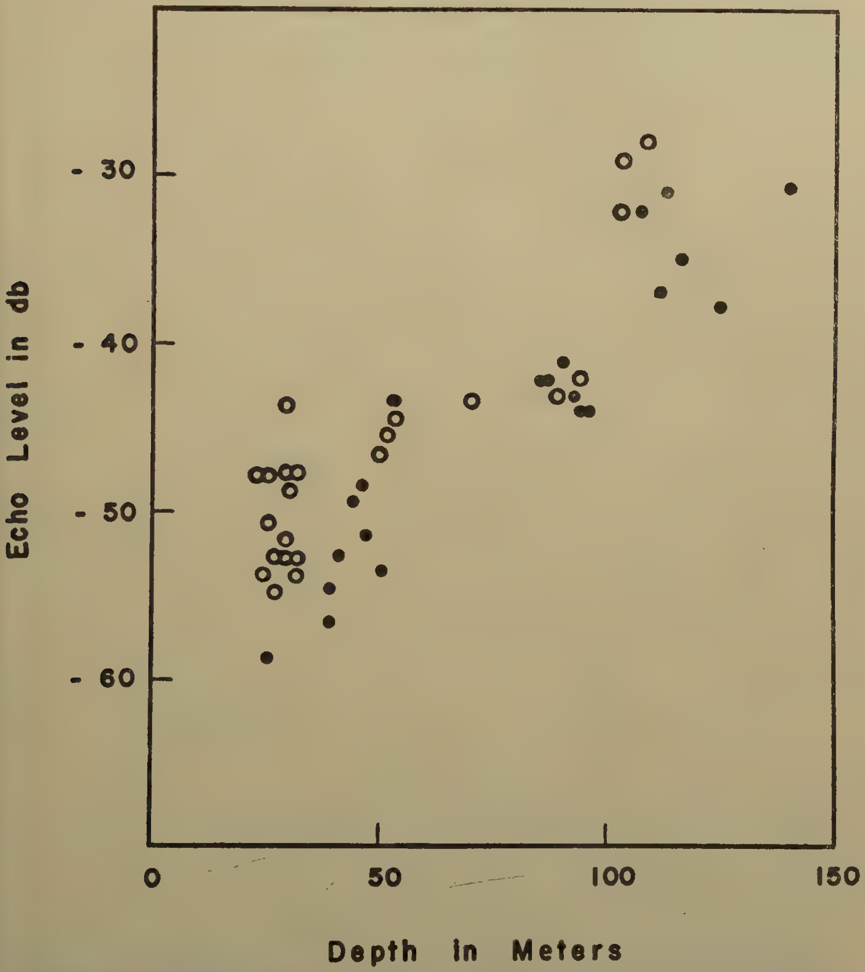


FIGURE 5. Echo levels recorded on March 17, 1945 at Lat. 22° 39′ N, Long. 109° 34′ W, at 1617 Pacific Standard Time.

the observed number of the larger scatterers increases with depth but does not necessarily show that the distribution function ν (σ) has changed significantly with depth. Even if this distribution did not change with depth, the increase in the number of echoes per unit range predicted by (5) and (6) could cause a depth effect of the type observed, provided the scatterers had a broad distribution of cross section.

A similar plot of the echoes from 11 oscillograms, recorded on March 17, 1945 at Lat. 22° 39′ N, Long. 109° 34′ W, at 1617 Pacific Standard Time, is shown in Fig. 5. Although fewer echoes were observed, the appearance of the plot is quite similar to Fig. 4.

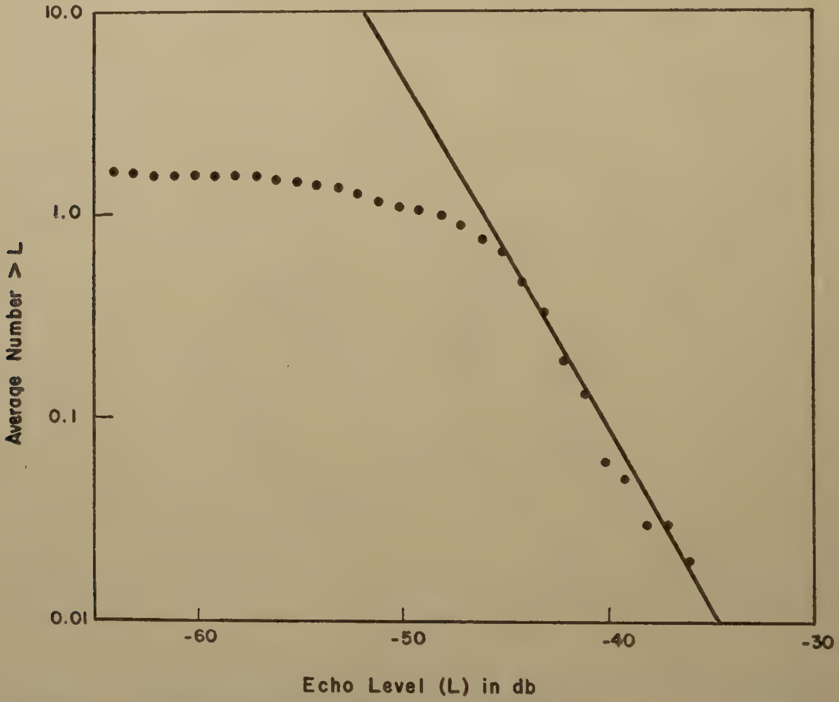


Figure 6. Cumulative distribution of echo levels for depths between 90 m and 180 m.

The results in a single set, such as those illustrated in Figs. 3 and 4, are characteristic of the particular configuration of scatterers at a given place and time, and, owing to the small number of echoes, are subject to large sampling errors. In order to obtain a sample of echoes characteristic of the whole area investigated, large enough to be sig-

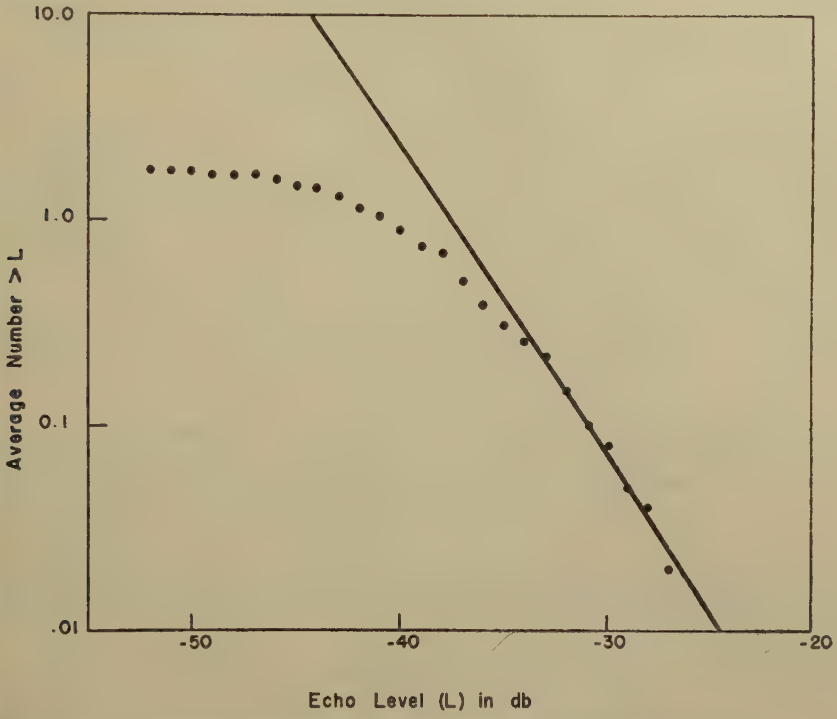


FIGURE 7. Cumulative distribution of echo levels for depths between 20 m and 90 m.

nificant, and at the same time unconfused by diurnal migrations of the scatterers (References *2*, *4*, *5*, *7*, *8*), quantitative study of echo distribution was made for the eight sets of records made during the daylight hours. Daylight was defined as the period between one hour after sunrise and one hour before sunset. Altogether 97 oscillograms were recorded in this group.

In order to study the effect of depth on the distribution, the echoes were divided into two depth intervals: 20 to 90 m and 90 to 180 m. The number of resolvable echoes deeper than 180 m was too small to be significant.

Cumulative distributions of echo level L for these depth intervals are shown in Figs. 6 and 7. The ordinates represent, on a logarithmic scale, the number of echoes per oscillogram with echo level greater than the values indicated in the abscissae. Observed frequencies are shown by the solid points. The solid straight lines were fitted to the upper ends of the observed distributions and represent the equations

$$\Phi\,(E) = 9 \cdot 10^{-6}\,E^{-1.75} \tag{8}$$

for Fig. 6, and

$$\Phi\,(E) = 2.4 \cdot 10^{-6}\,E^{-1.50} \tag{9}$$

for Fig. 7.

Echo distributions of this form are given by cross section distributions of the form

$$\nu\,(\sigma) = A\sigma^{-k}. \tag{10}$$

Substituting (10) into (6),

$$\Phi\,(E, R) = 4\pi R^2 A \int_E^1 \sigma^{-k} F\,(E/\sigma)\,d\sigma, \tag{11}$$

or, noting that $E/\sigma = \eta$,

$$\Phi\,(E, R) = 4\pi R^2 A E^{1-k} \int_0^1 \eta^{k-2} F\,(\eta)\,d\eta\,. \tag{12}$$

If $< A >$ is the average value of A between the depths R_1 and R_2, then the average number of echoes with intensity greater than E for the depth interval R_1 to R_2 is given by

$$\Phi\,(E) = \int_{R_1}^{R_2} \Phi\,(E, R)\,dR \tag{13}$$

$$= (4\pi/3)\,(R_2^3 - R_1^3) < A > E^{1-k} \int_0^1 \eta^{k-2} F\,(\eta)\,d\eta\,.$$

The integral in the right side of (13) is a function of k and is readily obtained by numerical integration. For the graphs of Figs. 6 and 7, respectively,

$$k = 2.75, \int_0^1 \eta^{k-2} F\,(\eta)\,d\eta = 0.00082 \text{ for Fig. 6,}$$

$$k = 2.50, \int_0^1 \eta^{k-2} F\,(\eta)\,d\eta = 0.0011 \text{ for Fig. 7.}$$

Solving (13) for $< A >$ by combining with (8) and (9):

$< A > = 3.6 \cdot 10^{-12}$ for depths from 20 to 90 m,

$< A > = 1.0 \cdot 10^{-10}$ for depths from 90 to 180 m.

It is of interest to examine the meaning of these figures in terms of the density of the scatterers. Let $< N(\sigma) >$ be the average number of scatterers per cubic meter with cross section greater than σ . Then

$$< N(\sigma) > = \int_{\sigma}^{\infty} < \nu(\sigma) > d\sigma \tag{14}$$

$$= < A > \sigma^{1-k}/(k-1).$$

Using the values of $< A >$ and k determined above, the cumulative distribution function $< N(\sigma) >$ is shown in Fig. 8 for both depth ranges, 20 to 90 m and 90 to 180 m, respectively. The solid lines

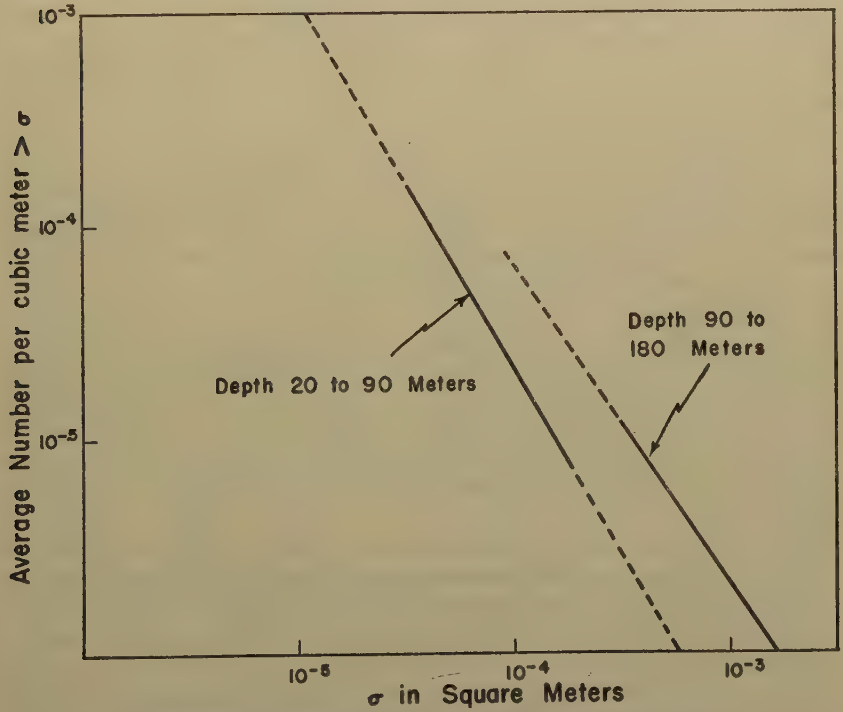


FIGURE 8. Cumulative distribution of scattering cross sections.

represent the portions corresponding to the most reliable portions of the echo distribution curves of Figs. 6 and 7. The dashed portions are less reliable extrapolations.

It is seen from Fig. 8 that the concentrations are quite small, being less than 10^{-4} per m^3 for cross sections greater than 10^{-4} m^2 (one cm^2).

Although the concentrations are greater for the 90 to 180 m depths than for the 20 to 90 m depths, the *form* of this distribution is nearly the same in the two cases. The somewhat smaller slope for the deeper scatterers indicates that the larger scatterers are relatively more frequent at the greater depths. However, this effect may not be significant, as it depends upon extrapolation of the distributions beyond the regions of most reliable measurement.

Volume Scattering Coefficients

Let $m\ (\sigma)$ be the average total cross section in square meters per cubic meter of all scatterers with cross section greater than σ . Thus,

$$m\ (\sigma) = \int_{\sigma}^{\infty} \sigma < \nu\ (\sigma) > d\sigma \tag{15}$$

and, for $< \nu\ (\sigma) > \ = \ < A > \sigma^{-k},\ k > 2$

$$m\ (\sigma) = \ < A > \sigma^{2-k}/(k-2). \tag{16}$$

Eq. (16) gives infinite values for m (0), a physical impossibility. This means that the distributions given by (14), and shown in Fig. 8, cannot be extended indefinitely towards smaller cross sections but must be limited by the fact that $m\ (\sigma)$ converges to a finite value for zero cross section.

This value, m (0), is the familiar volume scattering coefficient and represents the average total scattering cross section per unit volume for scatterers of all sizes. It has been measured at all of the locations of this study, using signals of 100 ms duration. The average values for the eight daytime locations used in the echo-distribution study were $3 \cdot 10^{-8}$ m^{-1} and $5 \cdot 10^{-8}$ m^{-1} at 90 m and 180 m depth, respectively.

In Fig. 9 plots of $m\ (\sigma)$ calculated from (16) for the 20 to 90 and 90 to 180 depth intervals are shown as solid lines. The dashed curves were sketched to indicate possible forms of $m\ (\sigma)$ which fit the calculated values and approach asymptotically to the measured volume scattering coefficients m (0).

Although this procedure is admittedly inaccurate owing to the lack of measurements of the smaller scatterers, Fig. 9 shows that, regardless

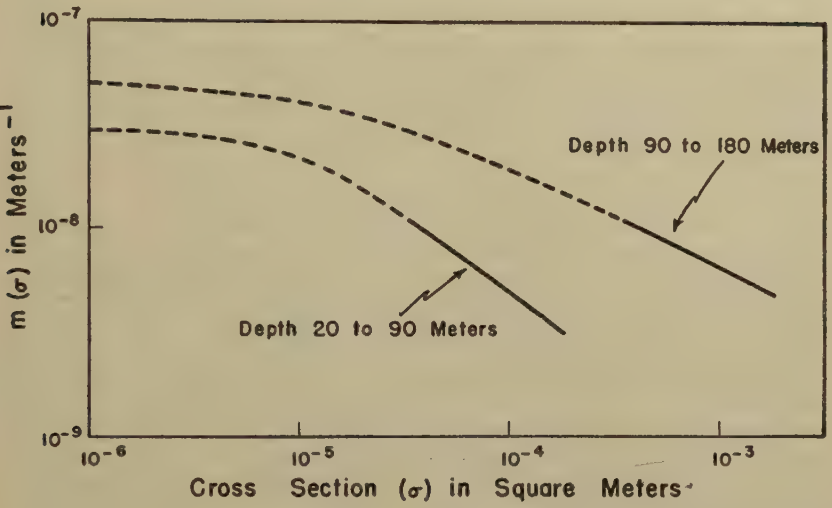


FIGURE 9. Volume scattering coefficient. Ordinates represent $m(\sigma)$, the total cross section in square meters per cubic meter of scatterers greater than σ. Solid lines represent values determined from the echo distributions. Dashed curves asymptotically approach the measured values of $m(0)$.

of the exact form of the dashed curves, most of the total observed scattering probably comes from scatterers with cross sections greater than 10^{-5} m^2 (0.1 cm^2).

DISCUSSION

The meaning of the above results in terms of the physical size, rather than the acoustic size, of the scatterers requires additional data on the reflectivity of the scatterers. Pending the acquisition of these data, one could speculate that, since the scatterers are probably biological organisms with density and compressibility close to that of water, their reflectivity will be low. Hence one would expect that the physical size would be appreciably greater than the acoustic size, particularly for the scatterers which are small in comparison with the wavelength of the sound.

However, the possibility of resonant scattering, caused by small gas bubbles in the scatterers or by other properties causing a large difference in compressibility with respect to water, cannot be excluded. If it were present to a significant extent, the physical cross section of many of the scatterers could be much smaller than their acoustic cross section.

Another factor deserving mention is the possibility of the grouping of several scatterers into colonies small enough to give the appearance of single echoes. If the thickness of a colony in the direction of the sound projector were less than one meter it might easily give the appearance of a single echo and be treated as a single scatterer in the analysis.

ACKNOWLEDGMENTS

The experimental records were obtained as part of a study of sea reverberation conducted under the joint direction of R. J. Christensen and the author by the Echo-Ranging Section of the University of California Division of War Research at San Diego. They were recorded by the sea-going group of the Echo-Ranging Section under the supervision of J. D. Frautschy. The excellent help in providing records of the high quality required in this study is gratefully acknowledged. The measurement of the oscillograms, and the computation of the results, was done by Miss Gloria J. Slack and Miss Gwendolyn A. Roy of this laboratory.

SUMMARY

Although the ultrasonic echoes from scatterers in the sea usually merge indistinguishably into the sound commonly called reverberation, discrete echoes are occasionally observed. Many such echoes, recorded at a frequency of 24 kc, were observed on a cruise conducted off Lower California in March and April 1945.

A study of the intensity distribution of these echoes has yielded estimates of the distribution of the scattering cross sections of the scatterers in the upper 180 m of the sea during the daylight hours. It was found that the concentration of scatterers decreased rapidly with increasing cross section. There were very few with cross sections greater than 10 cm^2. The average concentration of scatterers with cross sections greater than one cm^2 was somewhat less than 10^{-4} per m^3. Even though there were many more small scatterers than large ones, the very small ones did not contribute as much to the scattering as the larger ones. Most of the total scattering probably originated in scatterers with cross sections between 0.1 and 10 cm^2.

REFERENCES

1. ANONYMOUS
1943. Volume reverberation: Scattering and attenuation versus frequency. Univ. Calif., Div. War Res., Rep. No. U50. Publ. Bd. Rep. No. 39716.
2. 1946. Studies of the deep scattering layer. Univ. Calif., Div. War Res., File Rep. No. M445.
3. 1946. Forward scattering from the deep scattering layer. Univ. Calif., Div. War Res., File Rep. No. M398.
4. 1946. Stratification of sound scatterers in the ocean. Univ. Calif., Div. War Res., File Rep. No. 397. Publ. Bd. Rep. No. 60290.
5. DIETZ, R. S.
1948. Deep scattering layer in the Pacific and Antarctic Oceans. J. Mar. Res., *7* (3): 430-442.
6. EYRING, C. F., R. J. CHRISTENSEN AND R. W. RAITT
1948. Reverberation in the sea. J. acoust. Soc. Amer., *20:* 462-475.
7. HERSEY, J. B. AND H. B. MOORE
1948. Progress report on scattering layer observations in the Atlantic Ocean. Trans. Amer. geophys. Un., *29:* 341–354.
8. JOHNSON, M. W.
1948. Sound as a tool in marine ecology, from data on biological noises and the deep scattering layer. J. Mar. Res., *7* (3): 443-458.

UNDERWATER AMBIENT NOISE[1]

By

VERN O. KNUDSEN
University of California, Los Angeles, California
AND
R. S. ALFORD AND J. W. EMLING
Bell Telephone Laboratories, New York

INTRODUCTION

Underwater sound continues to be a problem of high priority in naval warfare. Since atomic bombing may render obsolete most types of surface ships, the relative if not absolute importance of the submarine certainly has increased. Thus, the U. S. Navy logistics chief recently told a congressional committee, "By far the most important and difficult problem which confronts the navy in so far as ship characteristics and fleet operating technique are concerned is the problem of undersea warfare."

Quite apart from its military applications, the purely scientific investigation of sound in the sea presents a fascinating "front" between the known and the unknown which invites further exploration. The scientific advances along this front that were incidental to the researches in subsurface warfare of World War II are of considerable significance, and it is almost entirely from these researches that we have gained our present knowledge of underwater acoustics.

The present report deals with a small but significant part of underwater acoustics; it surveys and compiles the principal available data (to March 15, 1944) on underwater ambient noise. It is condensed from a much longer survey report[2] which contains 76 pages of text and 146 charts and figures, prepared by the authors as a comprehensive reference work for naval and civilian groups concerned with underwater sound problems during World War II. The major part of the material of the original survey report was based on reports prepared by the following organizations: U. S. Naval Ordnance Laboratory; Massachusetts Institute of Technology—Bureau of Ships, U. S. Navy; Columbia University, Division of War Research; and University of California, Division of War Research. A complete list of reports consulted, together with a short summary of technical information applying to each, is given in the original survey.

---

[1] Based on a survey prepared by the authors under the direction of Section 6.1 of the National Defense Research Committee, Survey of Underwater Sound: Report No. 3—Ambient Noise, OSRD Report No. 4333, Sec. No. 6.1-NDRC-1848, September 26, 1944. The complete survey report has been declassified; photostatic or microfilm copies are procurable at the Publication Board, Scientific and Industrial Reports, Office of Technical Services, U. S. Department of Commerce, Washington, D. C. (The Pub. Board number of this report is 31021).

[2] See footnote 1.

TABLE I. General Locations at Which Ambient Noise Was Measured

Boston Harbor and off Massachusetts
Long Island Sound
Off Block Island
New York Harbor and approaches
Lower Chesapeake Bay and off Cape Henry
Off Beaufort, North Carolina
Off Florida and the Bahamas
Puget Sound
San Francisco Harbor and approaches
Off San Diego
Off Oahu and Midway Islands
Off Portsmouth, England
Loch Goil, Scotland
Off Hebrides Islands

Table I lists the areas in which the measurements of underwater noise were made. The data reported are confined largely to the frequency range of 100 c.p.s. to 25 kc. They are expressed in terms of pressure levels in db relative to 0.0002 dynes/cm^2. Often the results are given in terms of the "pressure level spectra," *i. e.*, the pressure level in a band one cycle wide, as a function of frequency. These values are derived, since the measurements are usually made with filters having a band width of about one octave, although in some measurements the band width is either less or greater than one octave.

QUALITATIVE NATURE OF AMBIENT NOISE

The term *underwater noise* is used to describe unwanted underwater sounds which tend to impair the operation of acoustically operated devices. *Ambient noise*, sometimes referred to as background noise, is the sound normally prevailing in water, usually from a multiplicity of sources such as water motion, marine life, and unwanted ship sounds. Usually it is not possible to specify quantitatively the contribution of sound from each of the sources present, but frequently a particular type of source is known to be preponderant. In such cases it is convenient to refer to the ambient noise more specifically as water noise, noise from marine life, ship noise, etc.

There are three main sources of ambient noise:

(1) *Water Motion*—Usually the prevailing source of noise in open and deep water. The magnitude is largely determined by the motion of the sea surface, particularly the number and prominence of breaking waves and whitecaps. Therefore, noise from this source is related to weather conditions.

(2) *Marine Life*—Produced by a large number of vertebrates and crustacea. Usually found in shallow tropical or semitropical waters. Diurnal and seasonal variations are common.

(3) *Ship and Man-made Sources*—Found in busy harbors and connecting waters. The magnitude depends on local conditions and may vary greatly from time to time.

Water noise is caused by a large number of widely distributed sources at the water surface; for this type of noise the field is believed to be essentially isotropic. An isotropic distribution is also to be expected in the center of large areas containing marine life. At a distance from such areas, or near localized concentrations of marine life, noise levels will vary with orientation and with the distance from the localized source. A similar situation will exist where the noise is produced primarily by a few near-by ships.

Variability is an outstanding characteristic of ambient noise. The magnitude of the noise from an individual source usually varies from moment to moment, and many sources exhibit large diurnal and seasonal variations. Some sources, such as ships, change their geographical position with time. Ambient noise is usually produced by many sources, some near the point of measurement and others located at a distance. Consequently, variations in the transmission loss in the medium are also a cause of variability. For these reasons ambient noise cannot be specified as a constant quantity but must be described in statistical terms. That is, an estimate can be made of the most probable (or average) amount of noise that is to be expected under given circumstances. In addition, it is frequently possible to estimate the degree of variability in terms of a frequency distribution *vs.* time, or in terms of the standard deviation when the distribution is known to follow the normal law. A knowledge of the average noise and the magnitude of the variation from the average is essential for the correct application of quantitative data on ambient noise.

NOISE FROM WATER MOTION

The term *water noise* is used to designate the underwater noise (at a sufficient distance from the shore to avoid the sound of waves breaking near the shore) produced by the agitation of the sea surface. This agitation is usually produced by wind. Water noise is the principal type of noise encountered in open and deep sea water. In practically all parts of the ocean, water noise is an important component of the noise, at least in some portion of the spectrum.

Noise is also produced by breaking surf, by the impact of rain and hail on the water surface, and occasionally by a movement of the bottom material, such as rock, gravel, shells, etc. Underwater springs, volcanoes, and gas vents also have been suggested as possible sources of noise. Surface ice and icebergs are subjected to stresses which often give rise to audible sounds. No satisfactory measurements of such noises have been made.

It is generally assumed that water noise is produced principally by breaking wave crests at the sea surface. It is believed that the unbroken undulations of the surface water do not contribute appreciably to underwater noise. The magnitude of water noise probably is related in a complex manner to such variables as height of waves, steepness of waves, and the number and magnitude of whitecaps present. Ideally, the estimation of water noise levels should be based on the combined effects of all these variables. Practically, this is not possible

TABLE II. METEOROLOGICAL SCALES

State of Sea			Beaufort Wind Force			
Scale No.	*Description*	*Height of Waves, Crest to Trough (Ft)*	*Scale No.*	*Description*	*Velocity m.p.h.*	*Velocity Knots*
0	Calm	0	0	Calm	<1	<1
1	Smooth	<1	1	Light Air	1–3	1–3
2	Slight	1–3	2	Light Breeze	4–7	4–6
3	Moderate	3–5	3	Gentle Breeze	8–12	7–10
4	Rough	5–8	4	Moderate Breeze	13–18	11–16
5	Very Rough	8–12	5	Fresh Breeze	19–24	17–21
6	High	12–20	6	Strong Breeze	25–31	22–27
7	Very High	20–40	7	Strong Wind	32–38	28–33
8	Precipitous	>40	8	Fresh Gale	39–46	34–40
			9	Strong Gale	47–54	41–47
			10	Whole Gale	55–63	48–55
			11	Storm	64–75	56–65
			12	Hurricane	>75	>65

with the data now available. Fortunately, however, it has been found that a reasonably good estimate of noise level due to water motion can be based on either wave height or wind velocity. A summary chart for making these estimates (see Fig. 4) will be presented after considering some selected but typical results of the noise from water motion.

Table II presents some meteorological scales which are in current use, and which are used in this report, for describing the state of the sea (in terms of wave height) and the Beaufort wind force (in terms of wind velocity in knots or m.p.h.).[3]

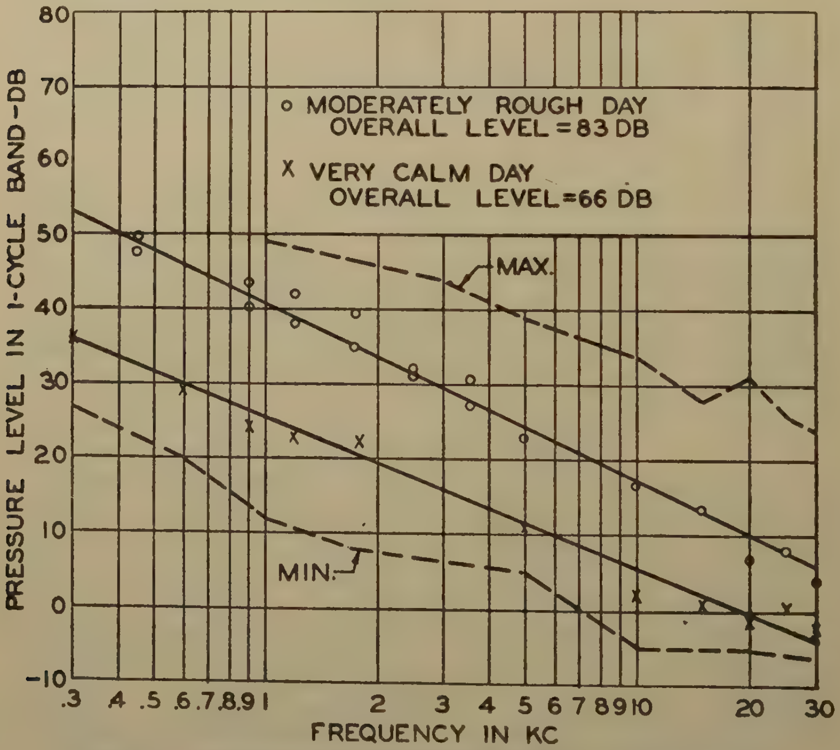


Figure 1. Pressure level spectra of ambient noise in Long Island Sound.

---

[3] In the original survey report, world charts prepared by the U. S. Weather Bureau are presented, one for each of the four seasons of the year, showing the distribution of average wind velocity, from which the average water noise can be estimated (with the use of Fig. 4). Thus, in one region of the North Atlantic, during December, January, and February, the average wind velocity is 24 knots, and, according to Fig. 4, the expected over-all (0.1-kc) noise level would be about 81 db.

Figs. 1–3 are illustrative of many results from measurements of noise, due mostly to water motion. Thus, Fig. 1 gives pressure level spectra measured in Long Island Sound on a calm day and on a moderately rough day. The over-all pressure levels (0.1–10 kc) corresponding to these spectra are 66 and 83 db, respectively, assuming a straight line extension of the curves to 100 c.p.s. On the same figure are also shown for comparison the maximum and minimum noise pressure levels measured from July to November 1942 in the waters adjacent to New York Harbor.

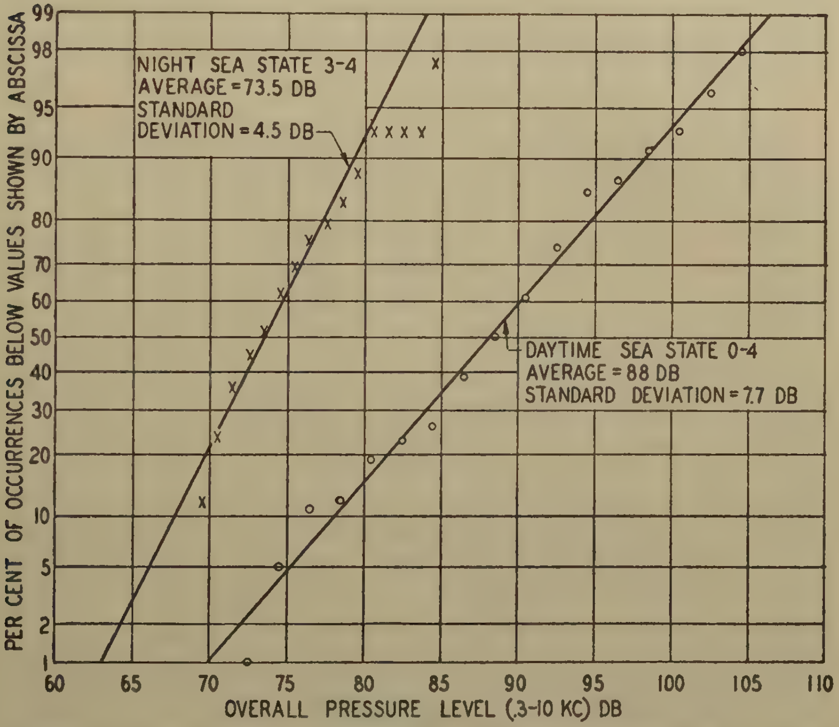


Figure 2. Ambient noise in approaches to New York Harbor; distributions of over-all pressure levels.

Fig. 2 is based on a noise survey of the area between Rockaway Beach and Sandy Hook conducted in March–April 1943. Pressure levels were measured throughout the frequency range of 150 c.p.s. to 25 kc, and over-all pressure levels were measured in a 0.3–10 kc band. Only the cumulative distribution of occurrence of over-all pressure levels for a series of daytime measurements (when ship sounds were predominant) and for a series of nighttime measurements (when ship sounds were almost completely absent) is shown in Fig. 2, plotted in the usual manner for determining the average value and the standard deviation. The curve for the nighttime measurements, when the sea state was estimated by the observers to be between 3 and 4, is believed to be free from man-made sounds, and probably therefore is representative of water noise. If the over-all pressure level measurements had been made with a band width of 0.1–10 kc instead of 0.3–10 kc, the pressure levels would have been about 2 db higher than those shown in Fig. 2. A comparison of the two curves of Fig. 2 reveals the marked difference between the daytime and nighttime underwater noise levels in the approaches to a busy harbor.

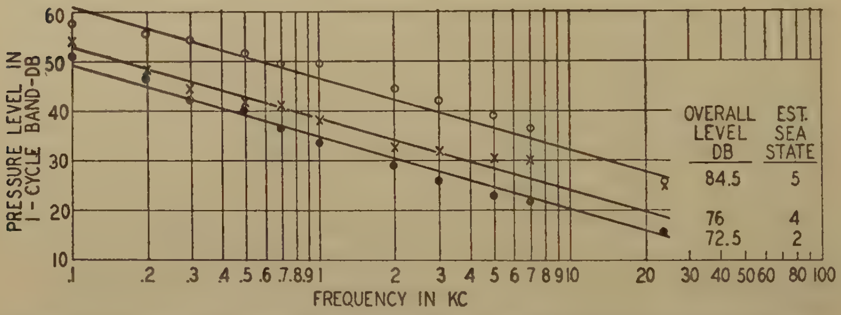


Figure 3. Pressure level spectra of ambient noise in open water about 700 feet deep.

Fig. 3 gives pressure level spectra of water noise in open and deep water five miles off Fort Lauderdale, Florida, for sea states 2, 4, and 5. The measurements were made by using a sensitive hydrophone having a low system noise. The hydrophone was carefully suspended from a cable by means of damped springs so that the movement of the hydrophone was negligible even in a very rough sea. The depth of submergence of the hydrophone was varied from 25 to 300 ft. No systematic variation of the over-all pressure level with hydrophone depth was observed; the small variations appeared to be random and probably resulted from the changing roughness of the sea and possibly from occasional sounds of distant ships or from marine life. Each point shown in Fig. 3 is the average of the data taken at all depths. There is a slight indication of a possible change in slope of the spectra for different states of sea, but the change is not systematic, and a reasonably good fit to the data is obtained with curves having the same slope as that of the spectrum determined by averaging all the data.

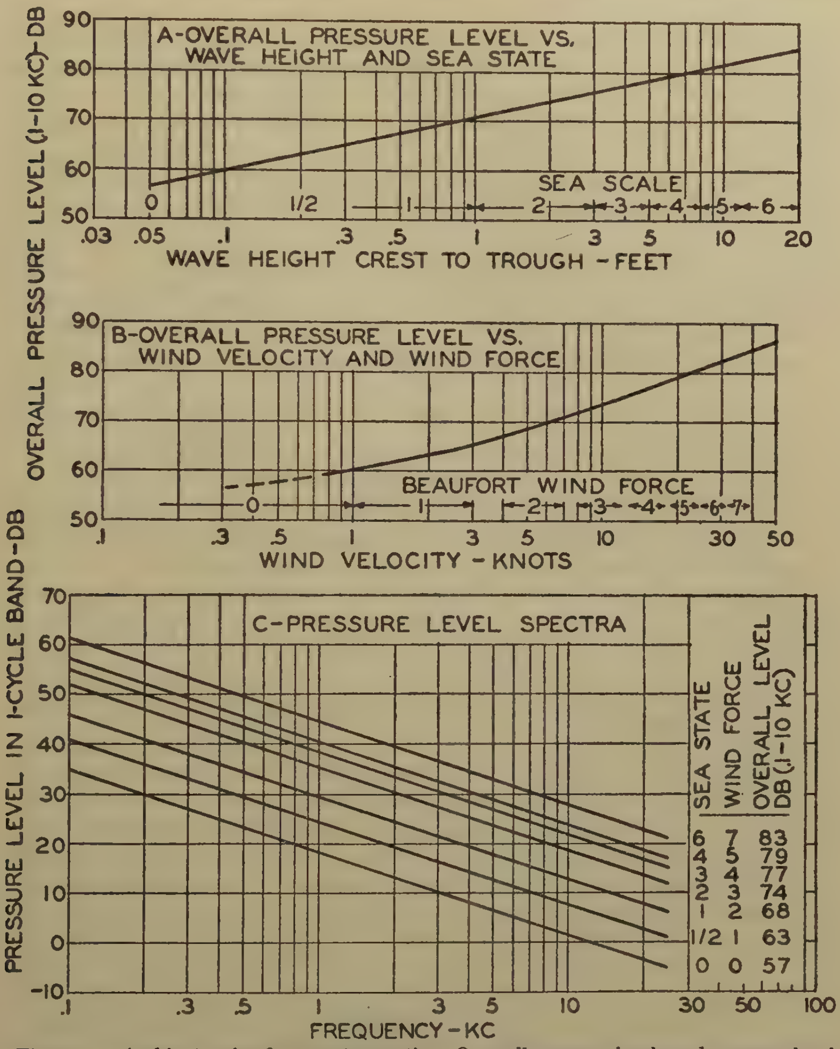


Figure 4. Ambient noise from water motion. Over-all pressure levels and pressure level spectra as a function of sea and wind conditions.

Fig. 4 presents composite summaries of ambient noise from water motion. The relation between over-all noise level and wave height (crest to trough) is shown in Part A. The "state of sea" scale corresponding to various wave heights is also shown along the abscissa. This sea scale provides a useful and an approximate method for describing the height of the waves.

The relation between over-all noise level and wind velocity is shown in Part B of Fig. 4. The Beaufort scale of wind force, which is frequently used for specifying approximate wind velocities, is also shown along the abscissa. The curves in Parts A and B are for average values. For a given wave height or wind velocity, considerable variation from the average noise level is to be expected. For example, noise levels higher than average are probable when whitecaps and spray are unusually prominent. On the other hand, noise levels lower than average are probable when the sea consists of long smooth crested waves. The standard deviation of observed noise levels with respect to the average values specified by the curves is of the order of 4 to 5 db. The variation is less at high wave heights and wind velocities than at low.

Frequently, when both wave height and wind velocity are known for a given location, the noise estimate based on wind and waves will differ. In such a case, an average of the two predicted noise levels usually will be more reliable than either one alone. An exception would occur near the lee shore of a large land mass. In this case, the estimate based on wave height will be the more accurate.

The average pressure level spectra to be expected for various wind and sea conditions are shown in Part C of Fig. 4. The slope of the spectrum appears to be independent of wind and sea and averages −5 db/octave. Experimental evidence indicates that random departures from this slope may occur, but usually the slope will not be more than −6 db or less than −4 db/octave. Neither noise level nor spectrum varies greatly with water depth so long as the water is sufficiently deep to prevent breaking of the waves. In deep water the average noise level and spectrum are essentially independent of the depth (20 to 300 ft.) at which the noise is measured. There is, however, a difference in the character of the noise. Near the surface, the noise from individual waves and whitecaps can be discerned, and the momentary variations in noise level are greater than at a greater depth.

NOISE FROM SURF, RAIN, HAIL, TIDE RIPS, AND MOVEMENT OF BOTTOM MATERIAL

The roar of breaking surf, which often sounds very loud to a submerged swimmer, is clearly indicated or registered on underwater listening or sound recording equipment. However, there are almost no quantitative data on the noise produced by surf. Observers at Cape Henry reported that during "rough weather" the over-all pressure level (0.1–10 kc) measured near the bottom, 300 yards offshore, was 78 db (corresponds to sea state 3–4).

The impact of rain and hail is believed to produce a noticeable amount of noise in the absence of other prominent noise sources. There are no data available on the effects of hail and only a small amount of data on the effects of rain. One set of measurements indicates that with a sea of state 1, or greater, the impact of rain has a negligible effect at frequencies below 1 kc but, owing to the presence of noise from marine life, the effect above 1 kc could not be determined quantitatively. Measurements in the Thames River at New London, Connecticut, showed an over-all pressure level (0.1–20 kc) of 75 to 76 db in the presence of a "steady but not torrential" rain as compared to 57 db prior to the rain. This agrees qualitatively with observations made by swimmers in lakes and rivers while they were swimming under water during summer showers of rain and hail.

One set of measurements made near a tide rip (current 3.5 knots) in a calm sea showed a rise in the pressure level spectrum beginning at about 3 kc. Levels of about 38 db (in a one cycle band) were obtained at frequencies of 5–10 kc (corresponds to sea state 8–9).

Sounds attributed to the movement of gravel on the bottom have been reported, but there appear to be no quantitative data on the noise from this source.

NOISE FROM MARINE LIFE[4]

Many forms of marine life are capable of producing underwater noise. Table III lists some pertinent information regarding the noise-producing characteristics of vertebrate specimens investigated at Shedd Aquarium, and of other soniferous specimens of marine life. Thus, the croaker, by means of drumming muscles on its air bladder, gives bursts of sound in rapid succession, each burst having a short duration and a principal frequency of about 250 c.p.s.; the porpoise "barks" and follows this with a "gobble, gobble" sound, similar to that of a turkey cock; the toadfish emits intermittent "boops"; the spot produces a series of raucous "honks"; snapping shrimp, by snapping closed an enlarged pincerlike claw, emit a "crackling" sound, not unlike that of burning dry twigs, or, as heard from a distance, like the sizzle of frying fat. Among fresh water soniferous fish, the giant boom, or singing catfish, when played on an angler's taut line, is re-

---

[4] Several others have reported results of measurements on noise from marine life. See D. P. Loye and D. A. Proudfoot, Underwater noise due to marine life, J. Acous. Soc. Amer., *18:* 446–449 (1946); F. A. Everest, R. W. Young, and M. W. Johnson, Acoustical characteristics of noise produced by snapping shrimp, J. Acous. Soc. Amer., *20:* 137–142 (1948); and M. W. Johnson, F. A. Everest, and R. W. Young, The role of snapping shrimp (Crangon and Synalpheus) in the production of underwater noise in the sea, Biol. Bull. Woods Hole, *93* (2): 122–138 (1947).

TABLE III. Noise-Producing Marine Life

Vertebrate Noise-Producing Specimens at Shedd Aquarium

Name	*Description of Sound*	*Sound Producing Mechanism*	*Principal Freq. CPS** *1*	*2*	*3*	*Length of Pulse Sec.**
Croaker	Bursts of drumming in rapid succession	Drumming muscles on air bladder	250	—	—	0.022
Spot-fin Croaker	Individual drum beats	Drumming muscles on air bladder				
Black Drum	Isolated groans	Drumming muscles on air bladder	50	80	150	0.270
Red Drum	Isolated groans	Drumming muscles on air bladder				
Garibaldi	Clicking, rasping	Pharyngeal teeth	7,400	1,000	150	0.011
Sea-anemone	Drumming, tapping	Pharyngeal teeth				
Single-striped Damozel	Drumming, tapping	Pharyngeal teeth	700	—	—	0.023
Coral-reef Damozel	Drumming, tapping	Pharyngeal teeth				
Common Triggerfish	Rasping, hissing, spitting	Pharyngeal teeth	700			
Surgeonfish	Sucking, rasping	Pharyngeal teeth				
Long-finned Pompano	Clicking, grinding	Hitting together dorsal fins				
Sheepshead	Crunching, grinding	Extrinsic food crushing	5,800	1,900	—	0.026
Razorfish	Low crunching	Extrinsic food crushing				
Hagfish	Loud crunching, grinding	Extrinsic food crushing				
Muttonfish	Crunching, sucking	Extrinsic food crushing				
Sting Ray	Crunching, grinding	Extrinsic food crushing				
Red Grouper	Clicking, grinding	Snapping teeth while eating				
Nassau Grouper	Clicking, grinding	Snapping teeth while eating				
Spadefish	Thump from rapid motion	Mechanical disturbance of water caused by quick rotation of body	220	150	—	0.150
French Angelfish	Thump from rapid motion					
Black Angelfish	Thump from rapid motion					
Diamond Flounder	Grating, scraping	Disturbance of gravel				
Sea Catfish	Popping, drumming	?				

Other Reported Noise Producers

Vertebrates *Name*	*Remarks*	*Crustacea* *Name*	*Remarks*
Porpoise	Bark and gobble (observed off Cape Henry)	Snapping Shrimp	Crackling (principally *Crangon* [*Alpheus*] and *Synalpheus*)
Toadfish	Intermittent "Boops" (Beaufort, N. C.)		
Foolfish			
Grunt		Crabs	Especially *Cancer* and *Portunus*. Noise only while feeding
Sand Perch			
Squeteague	Only males have drumming muscles	Barnacles	Occasional clicks of low intensity
Sargo		Mantis Shrimp	Sharp click
Midshipman		Spiny Lobster	Grating
Gray Trout	Probably like croaker		
Spot	Series of raucous honks (Beaufort, N. C.)		
Sea Robin	Squawk, squeal or cackle. Similar to croaker but more rasping (Beaufort, N. C.)		

* From photomicro analysis of phonograph recording grooves.
Note:—The list given above may not be complete, not all noise producers being of equal importance; several other crustacea produce noise.

puted to sing "deep purring music that wanders up and down four full tones." Two of the forms of marine life, the snapping shrimps (*Crangon* and *Synalpheus*) and the croaker, are of great practical importance because they produce sustained noises of high level.

Snapping Shrimp. Snapping shrimp can be expected throughout the oceans at locations where environmental conditions are favorable. These conditions are:

Temperature—Limited by the 52° F winter surface isotherm. Some period of the year with about 60° F also required.

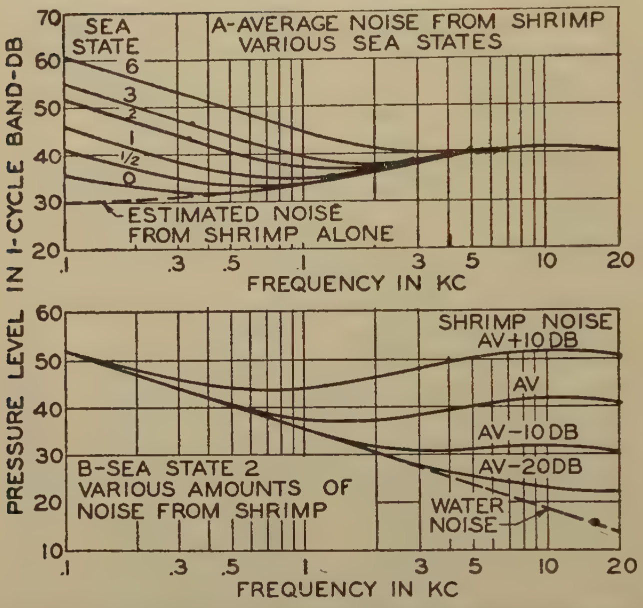
Depth—Generally less than 180 feet. The highest noise levels appear to occur in water between 30 and 150 feet deep.

Bottom—Rock, coral, shell, weed, or other material providing ready concealment. Relatively uncommon on mud or sand bottoms which are free from sheltering material.

The region in which shrimp are to be expected when depth and bottom conditions are favorable is roughly defined by a belt around the earth extending from latitudes 40° N to 40° S. Along the west coast of Europe extending north to Lands End, England, there is an additional area where the temperature is favorable. The inhabitable range along the west coast of South America south of Latitude 10° S and along the east coast south of Latitude 30° S is not known.

Since snapping shrimp are nonmigratory animals they can be considered as a constant characteristic of any region in which they have once been found. The noise produced by shrimp is continuous. There appears to be no pronounced seasonal variation and but slight diurnal variation. During the night hours, the noise level is usually a few db higher than during the daytime, and there is a peak in the noise level (3 to 4 db above daytime level) just before sunrise and another after sunset.

The sounds from snapping shrimp have been reported by others (ftn. 4, p. 419) and therefore they will not be considered further here, except for reference to Fig. 5 which gives summary curves typical of the ambient noise spectra to be expected in waters where snapping shrimp are prevalent. The data for this figure indicate that for frequencies up to about 1 or 2 kc the noise spectrum is largely determined by water noise. Above 2 kc the crackling of shrimp is the major source. Part A of Fig. 5 shows noise spectra with average shrimp noise and various sea conditions. Average shrimp noise is representative of the 24-hour average noise level found at water depths of 30 to 150 feet in the following regions: Off Florida, below Latitude 27° N, off Bahama and Cay Sal Banks, off San Diego, and off Oahu and Midway Islands.



OVERALL PRESSURE LEVELS (.1–10 KC)

SEA STATE	LEVEL DB
0	79.5
1/2	79.5
1	80.0
2	80.5
3	81.0
6	84.5

OVERALL PRESSURE LEVELS

SHRIMP NOISE	LEVEL DB
AV+10	89.5
AV	80.5
AV–10	75.0
AV–20	74.0

Figure 5. Pressure level spectra of ambient noise in the presence of snapping shrimp.

The noise levels at depths between 30 and 150 feet in these regions had a standard deviation of 5 db.

At depths greater than about 150 feet, shrimp crackle depends largely on the transmitted noise from near-by shallow water areas. At a distance of a nautical mile from such a shallow water area, the noise from shrimp is about 20 db lower than in the shallow water area. The colonies of shrimp appear to be smaller and more scattered at the colder margins of the areas which they inhabit. For example, along the eastern coast of the United States, shrimp have been found and their noise observed as far north as Latitude 35° N, but at offshore locations above Latitude 27° N the noise levels seem to be 15 to 20 db lower than the average level shown in Fig. 5. However, large colonies apparently occur in the harbors at Beaufort and Morehead City, N. C. Other isolated large colonies are probably to be found above Latitude 27° N.

The highest observed level of shrimp noise was measured near a pier at Kaneohe, Island of Oahu. The level at this location, at frequencies above 3 kc, was approximately 20 db above the average spectrum shown on Fig. 5. This is believed to be an exceptional condition caused by the presence of a large colony of shrimp living in the fouling material on the piling. Levels of this magnitude have not been observed in open water.

Spectra representative of a sea of state 2 and various amounts of noise from shrimp are shown in Part B of Fig. 5.

Croakers. Croakers (*Micropogon undulatus*), a variety of drumfish, are found during the late spring and early summer months in great numbers in Chesapeake Bay (estimated population 300 to 400 million). They also occur in considerable numbers at other east coast locations below Chesapeake Bay. The larger fish migrate to sea during the winter months. Different species occur on the west coast of the United States below Point Conception but apparently not in as great concentrations as along the eastern coast. Most of the available noise data have been obtained at east coast locations, and the discussion will be confined to the species found there.

Croakers produce noise by the contraction of drumming muscles attached to the air bladder. The sounds produced by an individual croaker consist of a series of "taps" that continue for about one and one-half seconds at a rate of about seven taps per second. The series is repeated at intervals of 3 to 7 seconds. The sounds resemble the tapping of a woodpecker on a dry pole. In an area with a high concentration of croakers, the sound during a period of great activity is a continuous roar, and the sounds of individuals may be heard only infrequently over the chorus of the whole population.

Croaker noise occurs principally during the feeding period, which starts in the evening as the bottom begins to darken. This, together with the migratory habits of the fish, accounts for the very pronounced diurnal and seasonal variation in croaker noise discussed in the following paragraph. Croakers are believed to feed in the shallow water on the slopes of banks.

Fig. 6 presents typical data on the seasonal and diurnal variations of over-all noise pressure levels from croakers at Cape Henry, Virginia. Note that the average over-all level in early June is 110 db, although levels as high as 119 db have been observed.

Typical spectra of ambient noise for various sea states in the presence of croakers (lower Chesapeake Bay) during hours of maximum activity are shown in Fig. 7 for late May to early June and for early July.

Other Marine Life. Other forms of marine life which may be of importance are toadfish, sea robins, some species of drumfish other than croakers, porpoises,[5] and a number of unidentified sources; one

---

[5] Some marine zoologists believe that the sounds here attributed to porpoises were produced by dolphins.

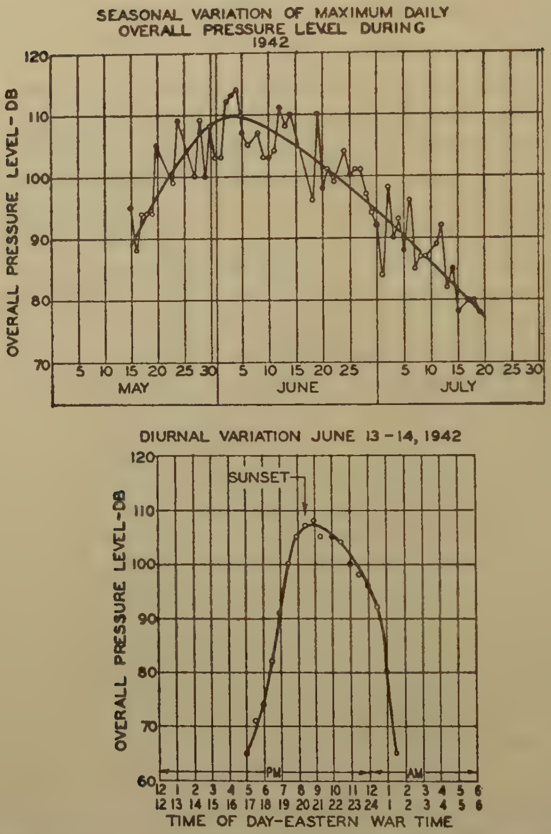


Figure 6. Seasonal and diurnal variation of over-all noise pressure levels from croakers.

such source is described as the sound of a mewing cat and another as an "awesome moaning."

Toadfish, individually, seem to produce higher noise levels than any other form of marine life thus far identified and reported, with the possible exception of the porpoise. The sound produced is an intermittent, low-pitched "boop" of about one-half second duration, similar to a boat whistle or sometimes like the cooing of a dove. The spectrum shown in Fig. 8 was measured very close to a single specimen. It will be appreciated that the levels decrease rapidly as the distance from the fish increases. These fish are shallow water bottom dwellers which nest under rocks, tin cans, and similar debris. Toadfish are not gregarious and apparently do not occur in sufficient concentration

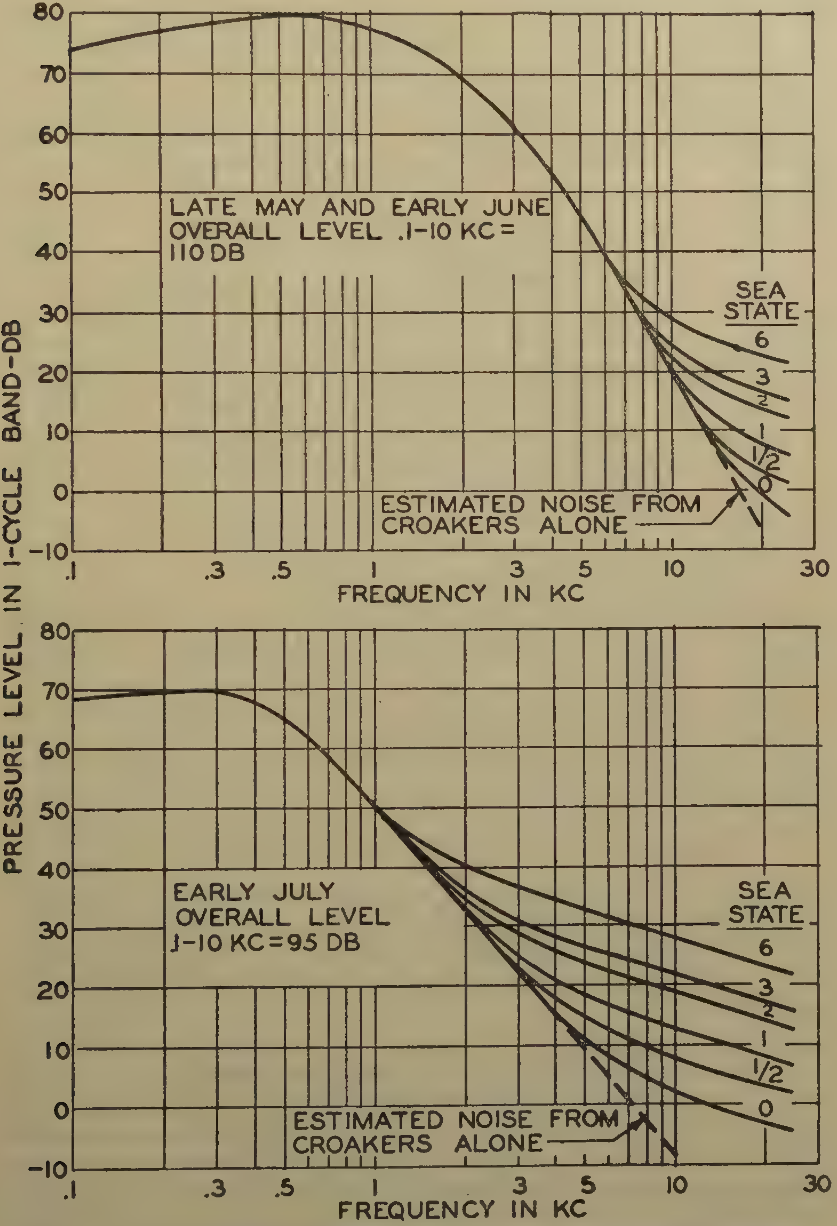


Figure 7. Typical prsssure level spectra of ambient noise in the presence of croakers.

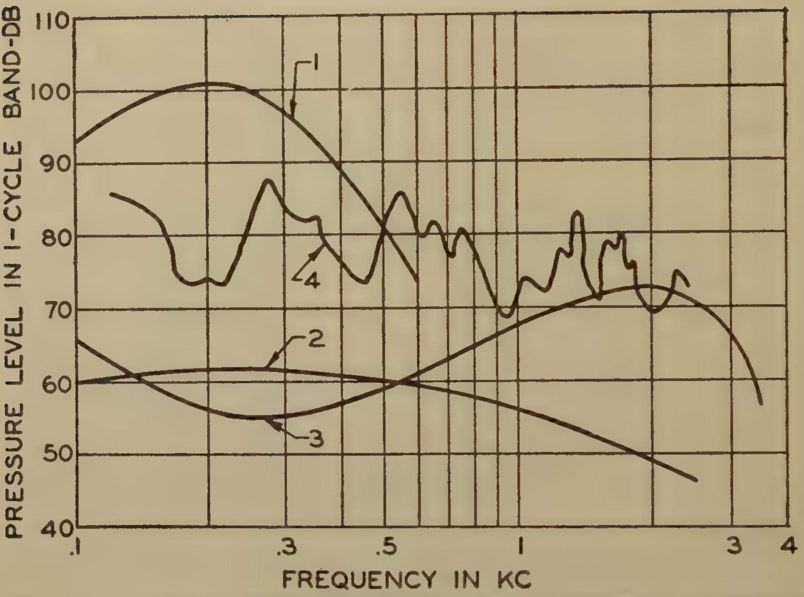


Figure 8. Illustrations of ambient noise spectra obtained in the presence of miscellaneous marine life. Curves: 1. Individual toadfish very close to hydrophone. 2. Principally from sea robins. 3. Unidentified "high pitched drumfish," possibly bastard trout. 4. Porpoises.

to produce the continuous roar generated by croakers during peak activity. Toadfish noise is subject to little diurnal variation. Other spectra shown in Fig. 8 are for (2) sea robins, (3) bastard trout, and (4) porpoises.

SHIP AND MAN-MADE SOURCES OF NOISE

In and near busy harbors and industrial centers, ships and manufacturing activities are the principal sources of ambient noise.

A wide variety of noise may be produced by industrial activities, and it is usually impossible to predict the noise from such sources in specific locations.

The noise produced by ships is more uniform in character than industrial noise, but it varies in magnitude over a wide range. It obviously depends on the number of ships present, their speed, their distance from the point of measurement, and the sound transmission characteristics of the water. Measurements in a number of east coast harbors showed an average over-all (0.1–10 kc) noise level of about 80 db. Location averages varied greatly, ranging from 65 db to 103 db. The standard deviation of the location averages was 8 db.

The average level of 80 db appears to be representative of locations with a moderate amount of ship traffic, such as the area near ship lanes at the upper end of Long Island Sound. An over-all pressure level (0.1–10 kc) of 90 db is more nearly representative of locations with a large amount of ship traffic, as, for example, in the approaches to New York Harbor during the daytime.

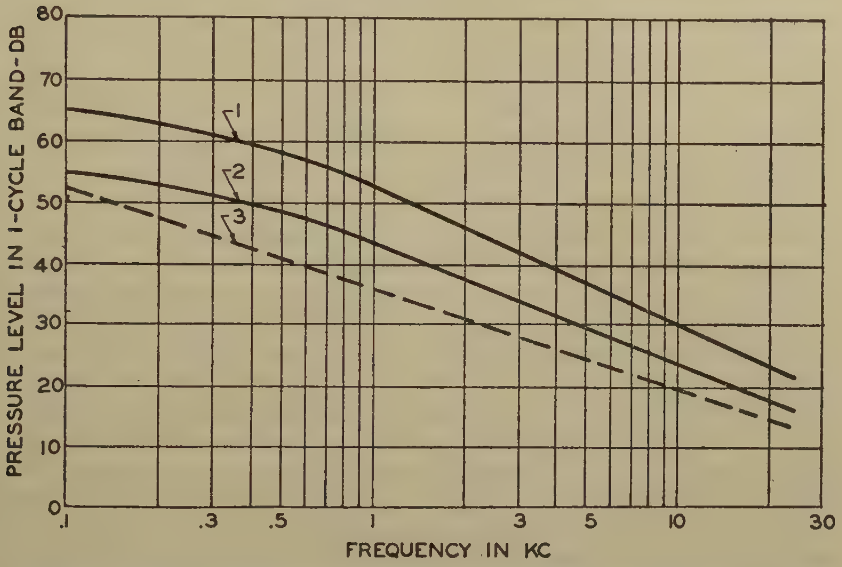


Figure 9. Pressure level spectra of ambient noise in the presence of ship sounds. Curves: 1. High noise locations; example, entrance to New York Harbor in daytime—over-all level 90 db. 2. Average locations; example, upper Long Island Sound near ship lanes—over-all level 80 db. 3. Water noise, sea state 2—over-all level 74 db.

Pressure level spectra representative of moderate and busy locations are shown in Fig. 9. A spectrum of water noise with a sea of state 2 is shown for comparison. In locations with light ship traffic, the average noise level is not likely to be lower than that produced by a sea of state 2.

PREDICTION OF NOISE CONDITIONS

The data compiled in the present report can be used for predicting the noise conditions to be expected in oceans, under various situations and in certain locations. Obviously, if survey data are available for a specific location they are preferable to estimates based on the general characteristics of the location. The principles which should be followed in predicting noise are outlined briefly as follows:

Open and Deep Water. Water noise is the major component. When sea state (wave height) and/or wind force are known, over-all pressure levels and spectra can be estimated from the data in Fig. 4. If, for a given wave height, the sea is unusually choppy with exceptionally prominent whitecaps and spray, the noise levels probably will be somewhat higher (4 to 8 db) than the average values shown in Fig. 4. If, on the other hand, the sea consists largely of long swells with exceptionally few whitecaps and little spray, the noise levels probably will be lower than average. When the prediction is to be made for a short time in the future, wind and sea conditions can be estimated from a series of weather maps. For prediction far in the future, the probable wind conditions can be estimated from U. S. Hydrographic Office Pilot Charts.

Shallow Coastal Waters. Noise from three sources is to be expected —water motion, marine life, and ships. Occasionally, surf and the movement of bottom material will contribute to the noise, but data for estimating the noise levels from these sources are not available. Water noise is estimated as outlined above. Noise from marine life is probable in tropical or semitropical waters. Snapping shrimp are to be expected at locations where the winter surface temperature is higher than about 52° F, where the water depth is less than 180 feet and the bottom consists of rock, coral, or similar material. Charts showing the nature of the bottom and water depth will be useful in estimating the prevalence of shrimp. Noise levels and spectra to be expected in the presence of shrimp are shown in Fig. 5. Croakers are a likely source of noise during the evening hours between late May and August in Atlantic coastal regions between Chesapeake Bay and Cape Canaveral. No doubt there are other regions where croaker noise is likely to be of local importance. The spectra shown in Fig. 7 are probably the best *a priori* estimates of noise conditions. Croaker noise is likely to be variable, and local surveys are advisable at locations where croakers are to be expected. The likelihood of ship sounds can be estimated from a knowledge of ship traffic. The data in Fig. 9 are probably representative of the underwater noise near active coastal ship lanes. During the absence of ships, water noise and marine life will, of course, determine the ambient noise level.

Bays, Rivers, and Harbors. The three main noise sources are possible contributors. The procedure for estimating noise conditions is the same as outlined above for coastal waters with these exceptions: In warm waters near piers with old piling, noise from shrimp may be 10 to 20 db above average. Recent dredging operations may reduce the shrimp population. In busy harbors ship noise probably will be the predominant source during the hours of shipping activity. Average noise conditions in harbors comparable with the approaches to New York City can be represented by the "High Noise" spectrum of Fig. 9. Local conditions such as industrial development at the shore or the presence of unusual man-made noise sources may necessitate a local noise survey.

The spectra of the ambient noises illustrated in Figs. 1 to 9 are typical of some of the noise levels that have been measured. They indicate the wide range and diversity of underwater ambient noise.

SUMMARY

The forgoing report presents the results of a survey of available data (to March 15, 1944) on underwater ambient noise at frequencies from about 100 c.p.s. to 25 kc. The data are presented in the form of pressure levels (for a wide band of frequencies or for a band one cycle wide) in db relative to 0.0002 dynes/cm^2. The principal sources of ambient noise are water motion, marine life, and ship sounds. In the absence of sounds from ships and marine life, the noise level in deep water is largely determined by the state of the sea; the over-all (0.1 to 10 kc) pressure level varies from about 57 db in a "calm sea" to about 83 db in a sea under a strong wind (about 35 m.p.h.). Similarly, the underwater noise from croakers may reach over-all levels as high as 119 db, and the sounds of ships in busy harbors produce underwater noises that often exceed 90 db. Data on the average pressure levels of the principal sources of ambient noise are summarized. The actual noise encountered may differ appreciably from the average values; standard deviations vary from about 4 to 8 db. Principles are deduced for predicting the noise levels that may be expected under usual conditions in specific locations.

DEEP SCATTERING LAYER IN THE PACIFIC AND ANTARCTIC OCEANS

By

ROBERT S. DIETZ

U. S. Navy Electronics Laboratory
San Diego 52, California

INTRODUCTION

During Operation HIGHJUMP (U. S. Navy Antarctic Development Project, 1947), the writer frequently noted the presence of a layer of deep scatterers on the fathogram of the USS Henderson. This layer partially scatters the outgoing sound signal of the recording echo sounder during daylight hours so that a reflection is recorded which has the appearance of a false bottom at various depths between 150 and 450 fathoms.

Subsequent examination of a fathogram across the Pacific, as well as one from Hawaii to the Arctic, also revealed the frequent display of this phenomenon. As these records greatly increase present knowledge of the geographic distribution, and as other new information was observed regarding this phenomenon, the results of an examination of these fathograms is presented in this paper.

ACKNOWLEDGMENTS

The writer is indebted to C. Eckart and R. W. Raitt of the University of California, to M. W. Johnson of the Scripps Institution of Oceanography, and to R. J. Christensen, E. C. LaFond, and D. W. Pritchard of the U. S. Navy Electronics Laboratory for critically reading the typescript. The fathograms from which the data are taken were obtained by the personnel of the USS Henderson and the USS Nereus. In connection with obtaining the fathograms, the assistance of Captain C. F. Bailey of the USS Henderson and of E. C. LaFond, who supervised the obtaining of the Arctic fathograms of the USS Nereus, is especially appreciated.

PREVIOUS INVESTIGATIONS

Previous observations (Anon., 1942, 1943, 1946a, 1946b, 1946c) with sonar gear since 1942 by workers at the University of California Division of War Research and later at the U. S. Navy Electronics Laboratory, the Scripps Institution of Oceanography, and the University of California Marine Physical Laboratory, revealed that sound scatterers are not uniformly distributed in the ocean, but that they exhibit a striking variation with depth. Frequently, at about 175 fathoms, an exceptionally well defined layer of scatterers was found that had at least ten times the scattering power of scatterers at shoaler depths.

It was noted that this Deep Scattering Layer[1] (D. S. L.) is principally a daytime phenomenon. In the morning, scatterers descended from near the surface to form the D. S. L. and, in the evening, they ascended toward the surface. Such a diurnal cycle strongly suggested that this phenomenon was due to migrating marine organisms rather than to a physical discontinuity in the water, such as temperature change. Also, a temperature-change boundary is completely unable to account for the intensity of the scattered sound. It is well known from net hauls that offshore zooplankton, in general, such as the copepods *Calanus finmarchicus* and *Metridia lucens*, pelagic prawns, euphausiids, and many other organisms exhibit marked negative phototropism and that they make daily vertical migrations from the surface to depths as great as 2,500 feet. Presumably, many types of zooplankton migrate to the surface at night to feed in the phytoplankton-rich surface layers and, in the morning, they descend to regions of darkness at great depths, possibly for safety from their predators. However, many marine zooplankton forms may be too small to scatter 18-kilocycle sound effectively. For this reason, it has been suggested that the sound scatterers are nekton, such as fish or squid, which follow and feed upon the zooplankton. Recently, Lyman (1948) has suggested that squid are the scattering organisms. In any case, it is probable that migrating zooplankton are at least indirectly responsible for the D. S. L.

There is comparatively little previous information regarding the geographical distribution of the D. S. L. It is known to be frequently present off the California and Lower California coasts, especially during the day. It has been reported occasionally from various spots in the Pacific. Recently, workers at Woods Hole Oceanographic Institution have reported (personal communication) the presence of this phenomenon in the Atlantic Ocean. Also, it is probable that some reports of uncharted shoals by ships at sea are false bottoms ascribable to the D. S. L.

---

[1] It has also been suggested that this layer of deep scatterers be called the ECR layer in recognition of the joint discovery by C. F. Eyring, R. J. Christensen, and R. W. Raitt in connection with underwater sound work at the University of California Division of War Research. Other workers who have made important contributions to the study of this phenomenon include C. Eckart, G. E. Duvall, R. Ely, and M. W. Johnson. M. W. Johnson first showed the apparently biological nature of the scattering layer. Most of the studies by these workers are in anonymous wartime reports.

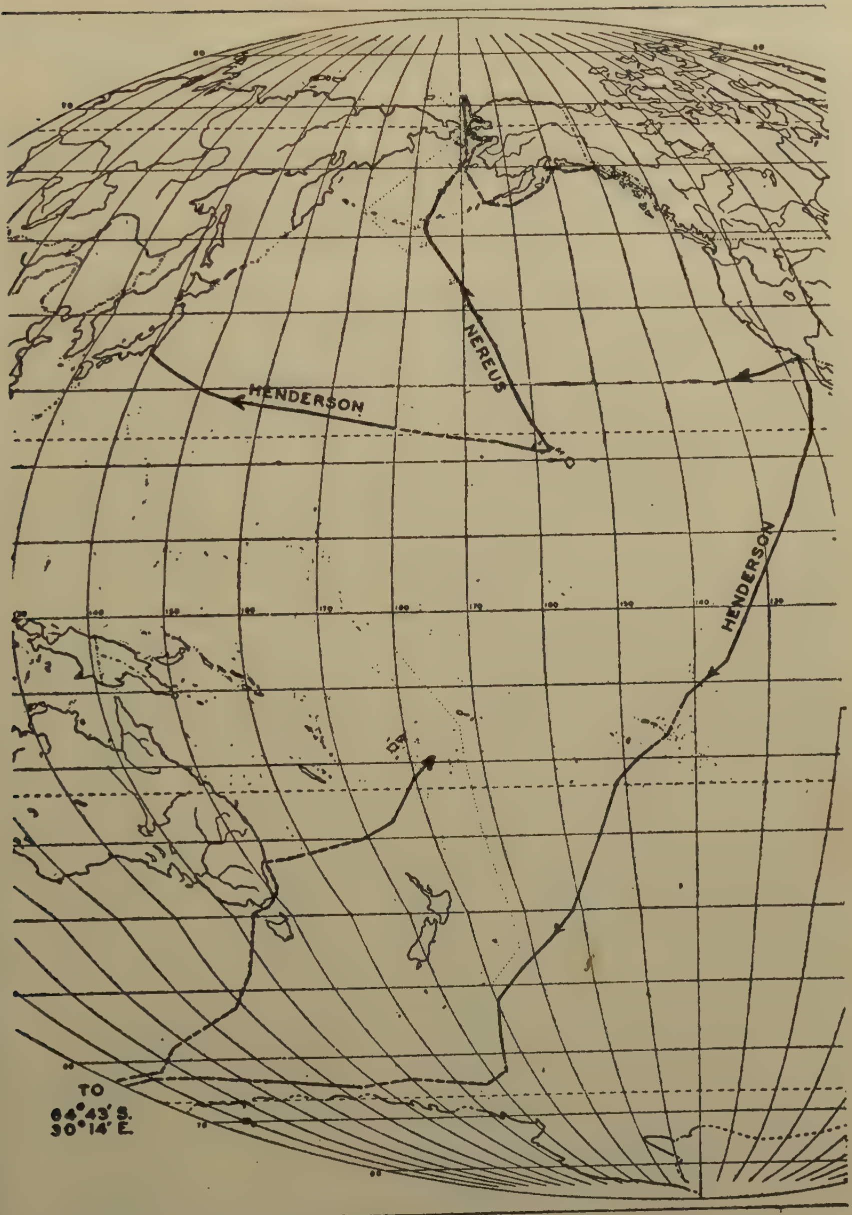
DEEP SCATTERING LAYER OF THE USS HENDERSON AND USS NEREUS FATHOGRAMS

General Remarks. Prior to a discussion of the HENDERSON and NEREUS tapes, it should be pointed out that the absence of a record of the D. S. L. on the fathogram does not necessarily preclude its presence, because such absence may be due to mistuning or other causes. A low gain setting will fail to bring in the D. S. L.; also, a gain setting which is too high may mask the Layer. Such changes in sensitivity undoubtedly account for its disappearance and reappearance throughout certain days. The Layer may also be masked by background noise produced by the ship when underway or by operation of some of the ship's equipment. However, when it does not show up throughout a long period of time it may be presumed to be absent or at least very poorly developed. Throughout these cruises, the echo sounders were operated primarily to obtain depth information, and the recording of the D. S. L. was incidental. The length of the outgoing sound signal masked any echoes from the D. S. L. when it ascended to depths less than about 100 fathoms on the HENDERSON record and less than about 60 fathoms on that of the NEREUS. The time used throughout this paper is the standard time for the zone in which the ship was located. The NMC echo sounder of the HENDERSON and the NMC-1 fathometer of the NEREUS utilized 18-kilocycle sound pulses.

Deep Scattering Layer during Operation HIGHJUMP. The distribution of the D. S. L. during Operation HIGHJUMP and other cruises is shown in Fig. 1, and its general appearance in Fig. 2. With the exception of two days in the vicinity of the Marquesas Islands, the Layer was recorded the greater part of each day during daylight hours throughout the passage from San Diego to the Antarctic between December 2–23, 1946. Controlled by the length of daylight, the Layer was detected at least discontinuously for as short as a ten-hour period and as long as a 19-hour period. Except in the early morning or the late evening, when the Layer was actively migrating, it was developed at depths from 150 to 450 fathoms. Near San Diego a depth of about 150 to 200 fathoms was characteristic, whereas in the tropical and southern waters a depth of 200 to 350 fathoms was more typical. The ascent and descent of the Layer was closely correlated with sunset and sunrise, respectively. With the exception of one or

---

Figure 1. Map of the Pacific Ocean showing the distribution of the Deep Scattering Layer along the tracks of the USS HENDERSON and the USS NEREUS. Solid lines indicate the continuous or almost continuous daytime development of the Layer. Dotted lines show the intervals during which indications of the D. S. L. were rare or entirely absent from the fathogram.


NEREUS
HENDERSON
HENDERSON
130
140
150
160
170
180
170
160
150
140
130
TO
64°43' S.
30°14' E.

two questionable sporadic displays of the D. S. L., it was completely absent from the fathogram during hours of darkness. The phenomenon of a double D. S. L. was frequently displayed.

The Antarctic Convergence was crossed at noon on December 23 in the vicinity of Latitude 62° S. This water mass boundary was clearly indicated by an abrupt drop in temperature from 39 to 33° F. It is noteworthy that the D. S. L. disappeared completely for three days after crossing this boundary and was rarely ever well developed in Antarctic waters.

During the period from December 25 to March 2, the Henderson navigated slowly westward along the western Pacific and Indian Ocean sectors of Antarctica. Only a few indications of layers of deep scatterers were recorded for short intervals from December 27 to January 21, which was a period of permanent (or almost permanent) daylight. From January 21 to 29 the D. S. L. was once again well developed and displayed the typical diurnal cycle (Fig. 2). This display might be correlated with the return of a day and night cycle, because by this date there were a few hours of complete darkness. However, subsequent to February 2 the D. S. L. was almost completely absent throughout the remainder of the Antarctic cruise and during the passage to Australia.

During the return passage from Australia to the United States, the D. S. L. was detected on the third day out of Sidney. The typical diurnal cycle was recorded each day until the echo sounder went out of operation at about Latitude 20° S. because of mechanical failure.

Deep Scattering Layer from Hawaii to the Arctic. During the period between July 15 and August 10, 1947, the USS Nereus made a continuous fathogram from Pearl Harbor to the Arctic with an NMC-1 echo sounder using 18-kilocycle sound pulses. This fathogram shows indications of the development of the D. S. L. each day throughout the cruise whenever the vessel was in deep oceanic water. The shoal depth of the Bering and Chukchi seas precluded the development of a deep layer of scatterers and the fathogram shows no evidence of any layers of scatterers in these areas.

The Nereus fathogram displays especially well the morning descent and the evening ascent of scatterers. They appear to migrate as a relatively compact layer rather than diffused, but the movement of scatterers in depths less than 60 fathoms cannot be determined because of masking by the outgoing ping. Although the diurnal vertical migration is well displayed, the scatterers descended to a depth so great that throughout most days they were beyond the range of good

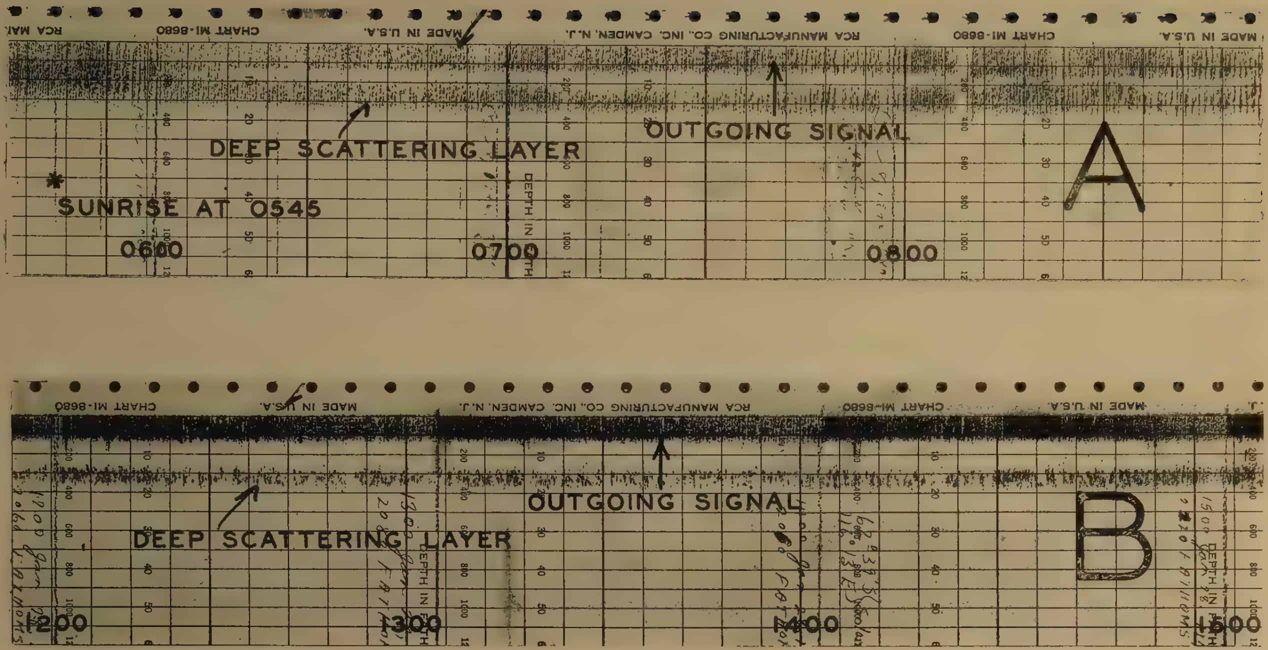


Figure 2

A. Fathogram showing the daytime development of the Deep Scattering Layer at 200 to 225 fathoms on December 8, 1946 in vicinity of Lat. 03° N. and Long. 131° W. Note the formation of the scatterers into a layer at about 0615 hours, or shortly after sunrise. Earlier downward descent of the scatterers is not apparent on this type of fathogram due to long length of the outgoing sound signal and general extension of this signal due to near-surface scattering.

B. Fathogram showing daytime development of the Deep Scattering Layer at 275 fathoms on January 28, 1947 in the Antarctic in the vicinity of Lat. 62° S. and Long. 117° W. This was one of the few instances in which the D.S.L. was well developed in the Antarctic.

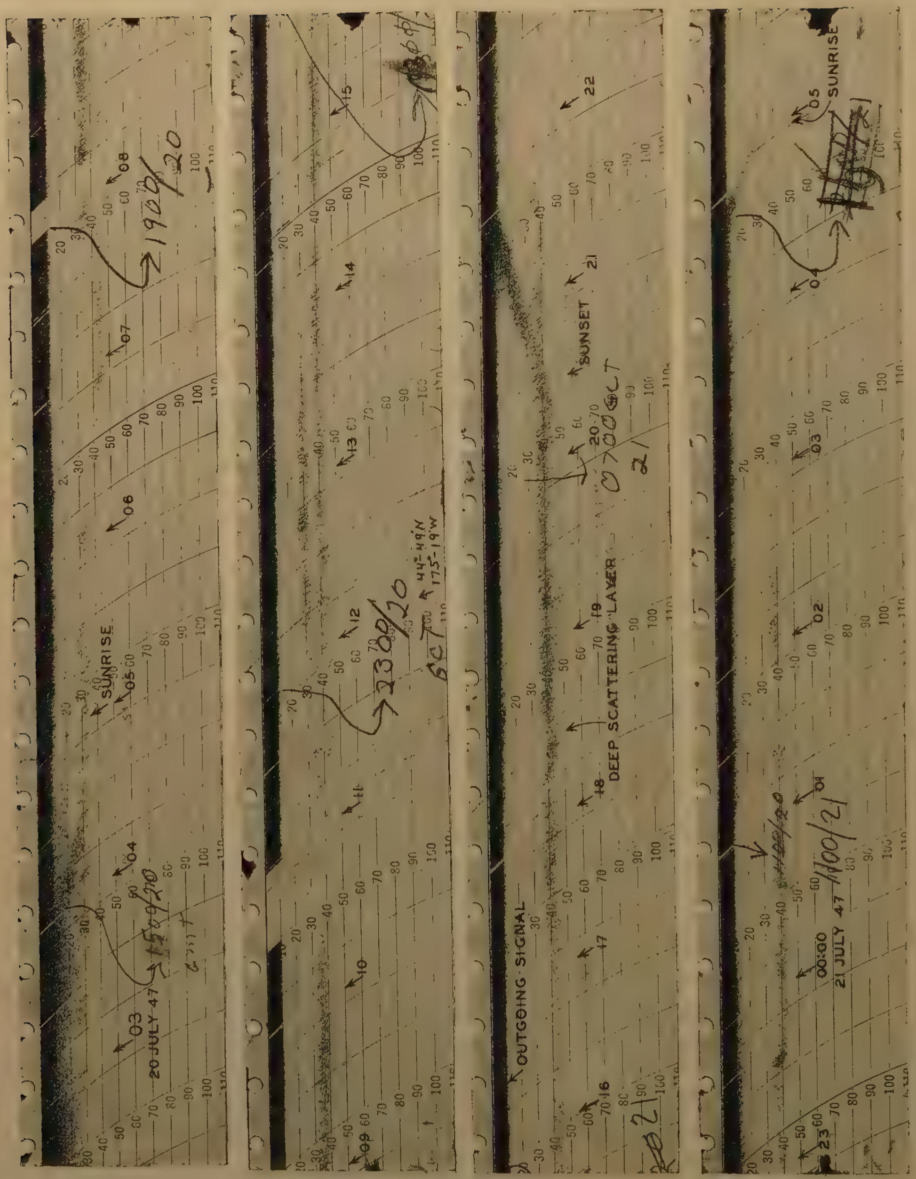

20 JULY 47
SUNRISE
SUNSET
DEEP SCATTERING LAYER
OUTGOING SIGNAL
00:00
21 JULY 47
44°-49'N
175°-19'W
1500/20
1900/20
2300/20
0700 OCT 21
1100/21

Figure 3

detection by the echo sounder; therefore the D. S. L. was only sporadically recorded.

On July 20 at noon position of Latitude 44° 49′ N., Longitude 175° 19′ W., the Layer was prominently developed throughout the entire day (Fig. 3). This comparatively strong development of the Layer may have been caused by a higher concentration of scattering organisms in the productive waters of the Aleutian (Subarctic) Current which lies between Latitude 42° N. and the Aleutian Islands. The development of the D. S. L. at depths from 175 to 225 fathoms was in agreement with depths observed elsewhere. Twice during this day, a double D. S. L. appeared. It was especially surprising that, although a portion of the scatterers rose at sunset, the rest remained at a depth of 175 fathoms throughout the night. Although the other tapes examined occasionally show sporadic indications of a scattering layer at night, this was the only record of a prominent and continuous D. S. L. throughout most of the night.

Another interesting phenomenon noted only on the Nereus fathogram was the presence of a double layer descending each morning during the period from July 15 to 18. These double descending layers were separated by a time interval of about 20 minutes and apparently were the result of the downward migrations of two types of organisms that are negatively phototropic to markedly different degrees. Yet, at sunset of these same days only a single ascending layer was noted.

The vertical migrations were so clearly defined on the Nereus tape that it was possible to measure roughly the rate of ascent and descent. In the morning the scatterers appear to descend at a rate of about 10 to 15 feet per minute. The evening ascent was slightly more rapid and was accomplished at a rate of about 15 to 18 feet per minute.

In agreement with all other observations, the D. S. L. was almost entirely a daytime phenomenon, and the vertical migrations were closely correlated with sunrise and sunset or, more exactly, with the beginning and end of a certain, but unknown, amount of twilight. The descent, deeper than 50 fathoms, began about one hour before sunrise, attained its maximum depth shortly after sunrise and varied only slightly throughout the day. The evening ascent began shortly before sunset but did not reach the 50-fathom depth until about one hour after sunset. Just what became of the layer at depths shoaler

---

Figure 3. Development of the Deep Scattering Layer in the North Pacific on July 20 and 21, 1947 at 150 to 175 fathoms (indicated depth scale should be increased by a factor of five). Note descent of the scatterers beginning before sunrise to form the D. S. L., and the ascent beginning shortly before sunset but not reaching the 60-fathom line until well after sunset. Note also the development of double Layers at 0700 and 1300 hours. Although most of the scatterers rise with nightfall, this fathogram shows the *unusual* development of a D. S. L. throughout the night.

than 50 fathoms is not known, because any further rise of the layer was masked by the outgoing signal. However, prominent extension of the outgoing signal during hours of darkness indicated the abundance of scatterers in the upper 50 fathoms of water.

Deep Scattering Layer from United States to Japan. During the period from August 11 to September 1, 1947, the USS Henderson obtained a fathogram across the Pacific Ocean with a recording NMC echo sounder from San Diego, California to Yokoshuka, Japan via Pearl Harbor, Hawaii. Although the echo sounder was operating erratically most of the time, resulting in a poor fathogram, the D. S. L. can be at least sporadically detected almost every day at various times during daylight hours. From the fathogram it appears probable that a properly working and sufficiently sensitive echo sounder would have recorded a daytime development of the D. S. L. essentially continuously across the Pacific.

When the vertical migration of the Layer was clearly recorded, it invariably began to descend about one-half hour or more before sunrise and to rise shortly before sunset; however, the upward migration was not completed until after sunset. Multiple layers were occasionally present and the scatterers were located at a depth of between 150 and 250 fathoms.

During the first half of the San Diego to Pearl Harbor passage, the D. S. L. was present during portions of each day at a depth of from 175 to 250 fathoms. Due to mechanical failure, the echo sounder was not in operation during the last half of this passage.

After leaving Pearl Harbor, the Layer was fairly well developed the first day out of port. During the next two days, it was not detected. Yet, from the fourth day out of port until arriving in the shallow water off Japan, five days later, the D. S. L. was recorded every day and it appeared to become more strongly developed with westward penetration.

DISCUSSION

Examination of three fathograms which cross the Pacific in both a north-south and an east-west direction, and cover a sector of the Antarctic Ocean as well, shows a wide development of the Deep Scattering Layer. It appears to be an oceanic phenomenon which is Pacific-wide and probably world-wide.

All of these records further substantiate the diurnal cycle by which the scatterers descend in the morning and ascend to near the surface at night. Judging from the Nereus fathograms, the scatterers begin their descent about one hour before sunrise and they do not complete their ascent until about one hour after sunset, so that they are in the surface waters only during rather complete darkness. Apparently the scatterers are motivated by a small amount of light, so that the migrations are actually correlated with a certain amount of twilight rather than with sunrise and sunset. No information was obtained on the distribution of scatterers in the surface water at night because of masking by the outgoing ping.

Because of the diurnal cycle, it is evident that the D. S. L. is a biological phenomenon so that the scattering agents must be zooplankton, nekton, or bubbles associated with such organisms. Yet, from these records no definite conclusion can be drawn as to whether zooplankton or fish cause the scattering. It will be necessary to make numerous laboratory experiments on the scattering characteristics of the various forms and to correlate underwater sound data with net hauls before the nature of the scatterers can be established. However, the extensive distribution of the D. S. L. is suggestive of zooplankton, and the comparatively slow speed of migration of the scatterers as a a layer is, perhaps, also more suggestive of the general zooplankton than of faster swimming fish. The occasional development of night layers, of double descending layers, and of multiple deep layers of scatterers show that the phenomenon is complicated and probably involves many kinds of organisms and/or different stages of development of a single species.

The amount of scattering is dependent in part on the sum of the cross sections of all the objects present. The total mass of zooplankton in the ocean is many times the mass of fish. Also, since the mass of a body varies as the cube of any linear dimension, whereas the surface varies as the square, a given mass of small organisms such as zooplankton present a much greater surface from which scattering might take place than an equal mass of fish. Thus, the extensive and frequently continuous distribution of the D. S. L. suggests that pelagic zooplankton may be the basic cause of the phenomenon. Yet, if zooplankton scatter sound like solid or liquid particles, most of them are too small to be effective scatterers of 18-kilocycle sound, which has a wave length of about three inches in water, because the scattering efficiency of an object smaller than the wave length is proportional to $(d / \lambda)^4$, where λ is the wave length of the sound signal and d is the circumference of the scattering object. Therefore, solid or liquid particles are inefficient scatterers if less than about one inch, and consequently an unreasonably high concentration of such organisms is required to produce the observed scattering. Also, the intensity of the scattering from small organisms is highly dependent upon the frequency of the sound, but such frequency dependence has been shown not to exist in work off California. Thus, although the small zooplankton such as the copepods are probably not the scatterers, larger forms such as the euphausiids, which often reach the length of about one inch, may possibly cause the scattering.

Very small air bubbles in water are known to be excellent scatterers of sound, because a gas has a markedly different density and compressibility than water. For this reason, gas bubbles have an "effective diameter" for scattering sound of many times their actual diameter. Scattering is especially great if the bubble is of a certain critical size so that it is resonant. The diameter of a resonant air bubble in water for 18-kilocycle sound is about 1.5 mm. at 150 fathoms, about 2.0 at 265 fathoms, and 2.5 at 415 fathoms. The size of a resonant gas bubble enclosed in a marine organism would vary somewhat from these figures because it would depend upon the elastic properties of the organism and the type of gas. In any event, if minute bubbles are enclosed in the migrating organisms, or if they are excreted by these organisms and exist for a short interval before being dissolved, they would probably adequately account for the phenomenon of deep scattering. However, there is no record of the presence of such bubbles in vertically migrating types of pelagic organisms.

As might be expected, there appears to be some correlation between the depth of the D. S. L. and factors which control the depth penetration of light. For example, previous observations (Anon., 1946a) in coastal water off California showed the characteristic depth to be about 150 to 250 fathoms. A similar depth was observed off California on the HENDERSON records. However, in the oceanic tropical and subtropical Pacific, under conditions of highly transparent water and an overhead sun, the D. S. L. was typically at a depth of from 250 to 350 fathoms. No good correlation was detected between the amount of cloud cover and the depth of the Layer.

An interesting question to consider is, What effect does the continuous daylight of polar midsummer have upon the diurnal cycle of the scatterers? From the HENDERSON fathogram it appears that the D. S. L. generally does not form under conditions of permanent or almost permanent daylight. Yet, weakly developed and questionable layers of deep scatterers were sporadically present. In this connection, net haul studies of plankton by Bogorov (1946) are noteworthy. According to his studies, zooplankton in the Barents Sea (Lat. 75° N.) under conditions of permanent daylight do not perform the regular vertical migrations that are characteristic of zooplankton of lower latitudes, but rather they maintain an almost unchanged vertical distribution throughout a 24-hour period. Further south in the White Sea (Lat. 65° N.) Bogorov noted the presence of a mixture of both the usual migrating types and the Polar non-migrating types of zooplankton.

Correlated with the return of a day-night cycle in the Antarctic towards the end of January, the diurnal migrating of the scatterers and the development of a D. S. L. was once again detected. However, in this connection the almost complete absence of the Layer during both February and the first part of March is puzzling and without a reasonable explanation.

Investigation of the D. S. L. is a fertile field for research. This phenomenon is of importance in connection with the transmission of underwater sound. If the scatterers are fish, even in part, a study of the Layer is of obvious direct commercial value. If the scatterers are zooplankton, much can be learned about these organisms concerning their habits and distribution. Also, if measurements are made under controlled conditions so that a scattering coefficient can be obtained. roughly quantitative data concerning the populations of the scatterers might be obtained. A study of the variation of the amount of scattering with sound pulses of various frequency would yield information on the size of the scatterers.

SUMMARY

Examination of almost continuous fathograms from California to the Antarctic, from Hawaii to the Arctic, and from California to Japan revealed a Pacific-wide distribution of the Deep Scattering Layer at various depths between 150 and 450 fathoms during daylight. As previous workers have noted, the D. S. L. is formed by scatterers that descend in the morning from near the surface and ascend in the evening. This diurnal cycle indicates that this phenomenon is produced by the vertical migrations of organisms rather than by a purely physical discontinuity. No direct evidence was obtained to determine whether the scatterers are zooplankton or larger forms such as fish. However, certain indirect evidence, such as the extensive and continuous distribution and the method of migration, is more suggestive of zooplankton.[2]

Although generally a daylight phenomenon, the D. S. L. was present throughout one night in the North Pacific. Double descending layers of scatterers were occasionally observed in the morning and multiple

---

[2] Since completing this paper and submitting it for publication, the writer has performed an experiment which shows that even very minute particles, when present in clouds, can strongly scatter high frequency sound. Sand clouds, made by pouring a bucket of fine sand (median diameter of 0.5 mm) into the ocean, have been observed to effectively scatter 18-kilocycle sound of an NMC-1 echo sounder down to at least 60 fathoms. Hence, it is likely that swarms of even the smallest zooplankton can account for the D. S. L.

deep layers are common, suggesting that many kinds, and/or developmental stages, of organisms are involved.

Controlled by the length of daylight, the D. S. L. was developed for as little as 10 hours per day and as much as 19 hours in high latitudes. Generally the D. S. L. did not appear to form during the period of permanent or almost permanent daylight of the Antarctic midsummer. Further investigation of this phenomenon promises to yield information of important scientific and commercial value.

REFERENCES

ANONYMOUS

1942. Reverberation studies at 24 kilocycles. Reverb. Group, Univ. Calif. Divn. War Res., Rep. No. U7.

1943. Volume reverberation: Scattering and frequency attenuation vs frequency. Reverb. Group, Univ. Calif. Divn. War Res., Rep. No. U50.

1946a. Stratification of sound scatterers in the ocean. Sonar Data Divn., Univ. Calif. Divn. War Res., Rep. No. M397.

1946b. Forward scattering from the deep scattering layer. Sonar Data Divn., Univ. Calif. Divn. War Res., Rep. No. M398.

1946c. Studies of the deep scattering layer. Sonar Data Divn., Univ. Calif. Divn. War Res., Rep. No. M445.

BOGOROV, B. G.

1946. Peculiarities of diurnal vertical migrations of zooplankton in Polar seas. J. Mar. Res., *6* (1): 25–32.

LYMAN, JOHN

1948. The sea's phantom bottom. Sci. Mon., N. Y., *66* (1): 87–88.

SOUND AS A TOOL IN MARINE ECOLOGY, FROM DATA ON BIOLOGICAL NOISES AND THE DEEP SCATTERING LAYER[1]

By

MARTIN W. JOHNSON

Scripps Institution of Oceanography[2]
University of California
La Jolla, California

In studies of specific populations or communities of different animals, the investigator is concerned largely with relative concentrations of individuals and with their feeding and breeding behavior, migratory habits, etc., as individuals or as groups. Information on these aspects is fundamental to an analysis and understanding of the factors that operate as controls within the population.

Direct census of population density based on systematic collecting, and identification or direct visual observations of the population *in situ*, are time-honored methods in ecological studies, and these will always be an essential part of such investigation whenever possible. Much tedious labor is involved, however, and therefore any physical means that provides a short-cut to useful generalizations or serves as a guide to focus attention on significant aspects, such as relative concentrations and activities in space and time, serves a real need. The application of sound and of underwater photography are two very promising means to this end. Both are still largely in the exploratory stages. The present discussion will be concerned only with sound.

In the past few years a large amount of work has been carried on pertaining to underwater noise phenomena, and a number of reports have been written relative to the behavior of sound in water, spectrum analysis, and identification of ambient noises, but it seems desirable to examine and interpret some of the information from an ecological rather than an acoustical point of view, to present new data and to point out some possible uses that suggest themselves.

---

[1] Most of the studies forming the basis of this report were made while the author was a member of the staff of the University of California Division of War Research at the U. S. Navy Radio and Sound Laboratory (now the U. S. Navy Electronics Laboratory) San Diego, California, under Contracts OEMsr-30 and NOBs-2074, respectively, with the Office of Scientific Research and Development (Section 6.1 NDRC) and the Bureau of Ships, Navy Department. Various members of the Listening and Echo Ranging Sections assisted in the studies, and acknowledgment is made especially to Mr. R. R. Carhart for preparation of Fig. 1, and to Mr. G. Duvall and Mr. R. L. Ely, physicists, who collaborated in study of the deep scattering layer.

[2] Contributions from the Scripps Institution of Oceanography, New Series, No. 380.

There are two principal ways in which sound, depending upon its origin, may serve in biological studies: (1) Underwater sound voluntarily produced by the animals and picked up on appropriate hydrophones and analytical equipment, and (2) sound projected as signals from sonar gear, such as from ships' fathometers or echo-ranging devices, and returned as reflected sounds by organisms.

UNDERWATER SOUNDS PRODUCED BY ANIMALS

The significant sound-producing marine animals fall into three general groups: (1) Crustacea, (2) fishes, and (3) mammals. Only the first two will be considered here.

Crustacean Noises

The crustacean sounds are by far the most widespread and persistent of all known underwater biological noises. This is due to the remarkable claw-snapping habits of two genera of snapping shrimps, *Crangon* and *Synalpheus*, widely distributed in tropical and subtropical waters (Johnson, 1943a).

The characteristic continuous loud crackle caused by populations of such shrimp is readily distinguished instrumentally from other ambient noises because of its continuity and high frequency (up to 50 kc) and great sound pressure spectrum level in the neighborhood of 10 to 15 kc.

The degree to which shrimp crackle dominates the noise in shrimp habitats has been discussed elsewhere (Johnson, Everest, and Young, 1947). It is instructive to consider here mainly the operation of directional hydrophones in seeking out these habitats. Fig. 1 shows results of a survey near San Diego when equipment of the type capable of 360 degrees rotation in a horizontal plane was used. The shrimp habitat at the point where these observations were made is a ledge of shale one mile in width in water of about 10 to 20 fathoms in depth. The shale is heavily pitted by borings and crevices made largely by boring mullusks.

It is seen from Fig. 1 that at a distance of 4,000 yards from the edge of the shrimp habitat the sound spectrum level at 20 kc is low and uniform in all directions, as shown by the solid black area in the center of the circles of pressure levels. Both depth and bottom conditions here were unfavorable to snapping shrimp, and the noise recorded is

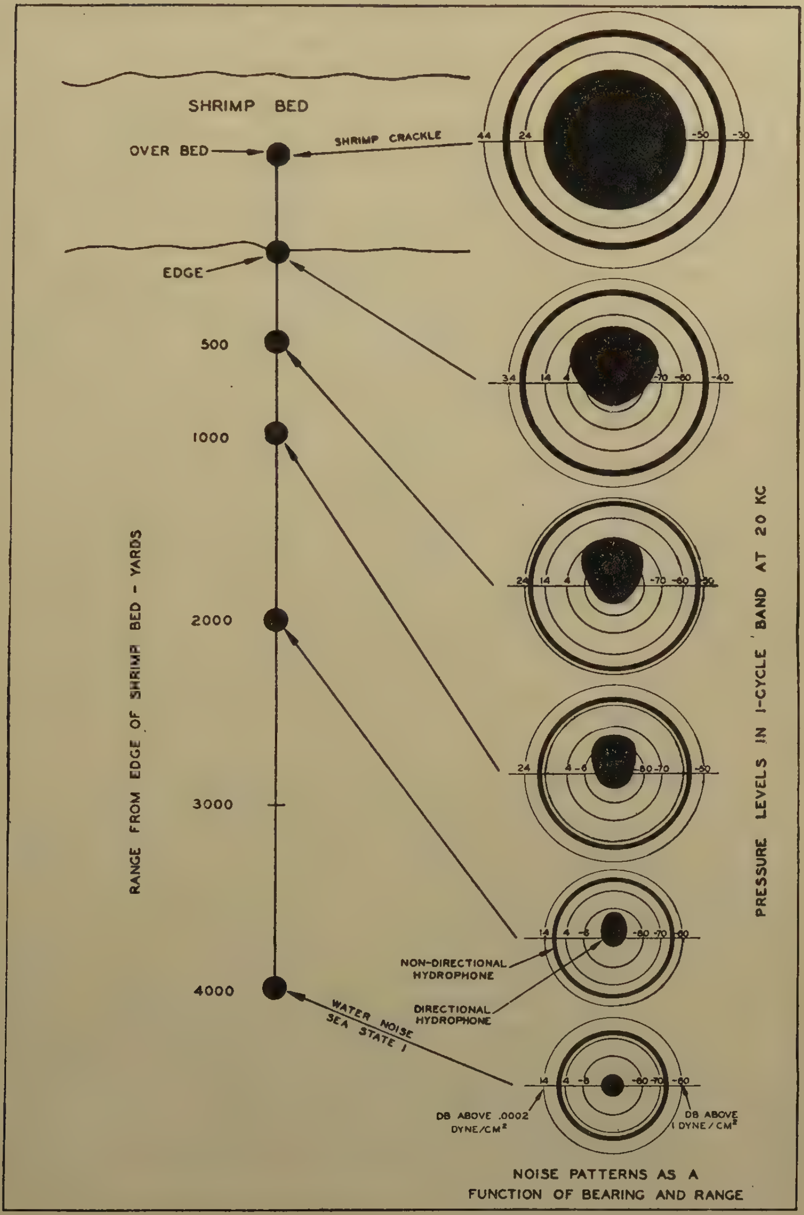

SHRIMP BED
OVER BED
SHRIMP CRACKLE
44
24
-50
-30
EDGE
500
34
14
4
-70
-60
-40
1000
24
14
4
-70
-60
-50
2000
24
4
-6
-80
-70
-50
3000
14
4
-6
-80
-70
-60
NON-DIRECTIONAL HYDROPHONE
DIRECTIONAL HYDROPHONE
4000
WATER NOISE
SEA STATE 1
14
4
-6
-80
-70
-60
DB ABOVE .0002 DYNE/CM²
DB ABOVE 1 DYNE/CM²
RANGE FROM EDGE OF SHRIMP BED - YARDS
PRESSURE LEVELS IN 1-CYCLE BAND AT 20 KC
NOISE PATTERNS AS A FUNCTION OF BEARING AND RANGE

Figure 1. Snapping-shrimp noise as detected with a directional hydrophone, with its beam rotated horizontally through 360°, over and at various distances from the shrimp bed.

"water noise." But upon approaching the ledge, there is a gradual increase of high frequency sound from the direction of the "shrimp ledge." Because of the directivity of the hydrophone, this increased sound spectrum level shows on the noise pattern as an expanding lobe in the direction of the shrimp bed. The angle of the high noise level increases as the bed is approached, indicating a habitat of great extent. When over the large shrimp bed, the sound spectrum level is again uniform in all directions, but because of intense shrimp crackle it is at a very much higher level than at 4,000 yards from the bed. Upon passing over the bed to the opposite side, where there is a shallow water area (not shown in Fig. 1) with sandy bottom, the sound spectrum level gradually diminishes and the lobe in the noise pattern again points towards the ledge, indicating that the ship now was over a habitat unfavorable to shrimp, apparently because of the sandy bottom.

The sound spectrum level obtained with the nondirectional hydrophone is shown as heavy circles in Fig. 1 for comparison.

In this manner additional information is gained besides that obtained from laborious dredging, which is often extremely difficult or impossible to accomplish. Illustrative of this are the results from dredging in an area of boulders at a depth of 10 fathoms off San Diego, where dredgings failed to yield more than an occasional shrimp, although the shrimp crackle was exceedingly loud over the area, indicating that shrimp are no doubt very abundant there (Fig. 2). Clean boulders have been brought up, but the agile shrimp appear to escape the slow moving dredge, so that their capture is only a matter of chance.

Direct and Indirect Uses of Shrimp Crackle

The crackling of shrimp may be used indirectly to some degree in the study of other benthic animals with which the shrimp are symbiotic or form more or less intimate environmental associations.

The possibility of using shrimp crackle in the study of certain large sponges, including the commercial sponge, should not be overlooked. It is well known that a number of snapping shrimp species, living within the pores and channels of sponges, can be heard when the sponge is removed from the water. A single sponge (*Spirastrella*) has yielded several hundred shrimp (Hay and Shore, 1915–1916). Though not yet demonstrated directly, there is reason to believe, on the basis of what has been said, that shrimp crackle might serve a commercial purpose in determining the location and extent of sponge concentrations.

The large amount of data showing a sharp drop in shrimp crackle at depths beginning at about 30 fathoms is convincing corroborative evi-

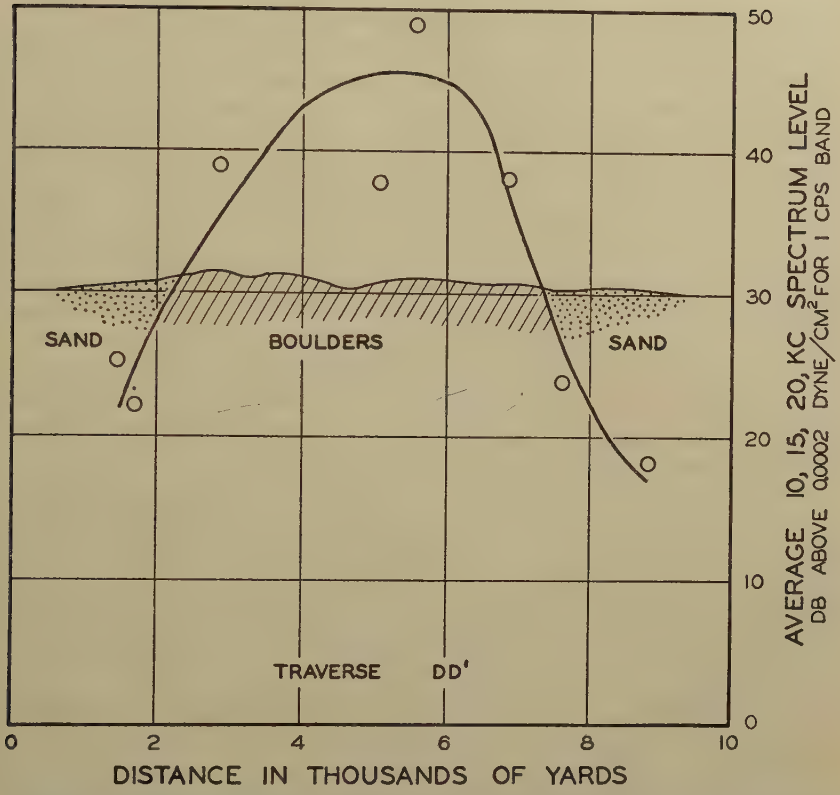


Figure 2. Sound spectrum level of shrimp crackle across a boulder patch in ten fathoms (from Johnson, Everest, and Young, 1947).

dence that this depth marks the maximum at which most species of *Crangon* and *Synalpheus* live in concentrations sufficient to provide instrumental measurements of their crackling. This is in keeping with an analysis of the vertical distribution of these genera as shown in 3,483 dredge hauls from various depths by several major expeditions, but the ease with which the information was obtained through acoustic means is in striking contrast.

The regularity with which shrimp crackle is found in, and is confined to, areas of certain bottom types, such as rock, coral, coarse shell, weed, and litter of harbors, is evidence of the importance that must be attached to the ecological factor which provides easily-maintained shelter for these animals.

If high sound level of crackle is an indication of the general type of bottom, and since other animals, including some fish, frequent (in some cases confined to) rocky or coral bottoms where snapping shrimp abound, then the noise may serve as a guide to these areas. Examples of such areas are the fishing grounds off North Carolina where isolated coral patches occur (Radcliffe, 1914). Shrimp crackle has been reported from one of these grounds by Dobrin (personal communication), and doubtless it occurs on others as well. On the California Coast, where the agariferous seaweed *Gelidium* is harvested on rocky bottoms, the divers have remarked that they can "hear the *Gelidium* crackle." Though this is an erroneous interpretation of the origin of the sound, the implication is clear.

Ecological Significance of Diurnal and Seasonal Variations in Shrimp Crackle

Certain ecological facts may be gleaned from a study of the diurnal and seasonal variations in shrimp noise.

Diurnal Variations in Activity. It is well known that many intertidal animals that can be observed during low tide, or at the water's edge, become increasingly active at night. Among these are the crabs, isopods, limpets, and other gastropod mollusks. This increased activity may be explained as an adaptation that is advantageous to the animals in avoiding daytime dessication during low tide and in avoiding the ravages of preying birds.

Is this instinctive nocturnal activity reflected equally in the animals that live in deeper water below the low tide level? A commensurate answer to this question is not readily obtained.

Acoustical measurements of shrimp crackle do, however, give evidence of a marked nighttime increase in noise. These measurements have shown a 2 to 5 db increase in sound pressure spectrum level of crackle during the night, with peak levels just after sunset and shortly before sunrise. The figures are logarithmic expressions, which, interpreted into terms of shrimp activity, reveal a two to threefold increase and perhaps reflect an over-all activity within the shrimp community. Involved in this activity are not only the shrimp but also other animals, who, by their increased wandering about, cause increased irritation to the shrimp and hence more frequent individual snapping.

Annual Variations. Studies on seasonal variations have been made only in the San Diego region where the seasonal variation in temperature is about 8° C. However, these studies do not reveal fluctuations in shrimp noise that are greater than those observed diurnally. From this it might be concluded that there is a rather uniform year-round population of adults capable of snapping, which implies that there may be a production of overlapping generations that do not migrate. In this connection it would be of interest to note what effect a greater range of temperature would have on activity in such an area as Beaufort, North Carolina, where the range is about 20° C, where snapping shrimp are abundant and where sound level due to them is known to be high during early spring and summer. A marked seasonal change in noise level might be inferred as a measure of the effect of temperature on the activity of the shrimp community.

Regarding the application of other crustacean sounds, little can be said. Several species of spiny lobsters produce a strident grating sound when they rub specialized structures on the bases of each of the antennae against opposing structures on the carapace. The sound is readily heard under water over acoustic instruments, but no attempt has been made to use it in locating populations. Indications thus far show that the sound is too rarely produced to be of general ecological use, but seasonal variations and local differences have not been sufficiently investigated.

Fish Noises

Many fishes are known to produce noises when removed from the water, but it is uncertain as to how many of these are soniferous in their natural habitat (Burkenroad, 1931). However, there are a number of notable sound-producers with organs well adapted to generation of underwater sound. Among these are especially the Sciaenidae (croakers and drums), the toadfish *Opsanus*, the singing fish *Porichthys*, and the searobin *Prionotus*. Additional underwater sound-producers are listed by Knudsen, Alford and Emling (1948), who also discuss the characteristics of fish sounds as related to "water noise." Other spectral analyses of fish noises are given by Loye and Proudfoot (1946) and Dobrin (1947).

The data available on fish sounds have not been collected or examined as closely with a view to their ecological application as was done for the shrimp studies. The most challenging suggestions come from the acoustical studies of the croaker, *Micropogon undulatus*, in Chesapeake Bay; the toadfish, *Opsanus tau*, at Beaufort, North Carolina; and various croakers at La Jolla, California.

The chorus of sounds produced by *Micropogon* in Chesapeake Bay was reported upon by Loye and Proudfoot (1946). Their findings include several items of ecological interest. (1) The chorus was heard only in the evening and nighttime. The interpretation suggested was that the noise accompanied a period of feeding. (2) The maximum nighttime intensities of the noise were localized to a specific area. (3)

Fish of this type appear to become accustomed to unusual noises, such as underwater explosions of blasting caps when these are detonated during the chorus. (4) Spectrum analyses of croaker noise in May 1942 showed appreciably more energy at frequencies up to 2,000 to 2,500 cps than three months later when there was very little above about 1,000 cps. It was suggested that the fish in the bay had grown during this period and that the correspondingly larger air bladder, which acts as a resonator, resulted in lower frequencies. However, the growth rate of sexually mature *Micropogon undulatus* is only about two to three inches in a year, according to Hildebrand and Cable (1930), so the size increase during three months could not be very great. Actually, the spectrum analysis probably revealed a means of testing a local population shift, which in this case resulted in the younger (smaller) fish being replaced in the area by an older population or by an admixture of older forms sufficient to lower the average sound frequency.

During a preliminary survey (Johnson, 1943b)[3] of underwater noise at Beaufort, North Carolina in April 1943, the most conspicuous noise encountered in Beaufort Harbor, other than shrimp crackle, was a soft cooing sound that was produced as isolated notes both day and night. This sound, later studied in detail by Dobrin (1947), was believed by him to be produced by the toadfish. The outbursts, as heard over the nondirectional hydrophone, always appeared to come from fixed locations on the bottom where they persisted for days at a time. They were found to be associated with the nest building habit of *Opsanus*. This is interesting evidence of persistent guarding of the nest or area as well as maintenance of a system of territorial rights, an instinctive habit known to be common also for some other fish.

In the summer of 1942 a chorus of fish noises was discovered at the outer end of Scripps Institution pier, which extends 1,000 feet outward from shore. Since that time, except for the spring and summer of 1946, and the summer of 1947, listening observations have been carried on at irregular intervals throughout the year. From these it was found that the chorus began in the spring about May 3, 1943; April 11, 1944; April 25, 1945; May 14, 1947. It continues throughout the summer and disappears about the middle or latter part of September. There are only occasional individual outbursts of sound at other seasons.

The chorus begins about sunset and increases gradually to a steady uproar of harsh froggy croaks, with a background of soft drumming.

---

[3] Thanks are due to Dr. H. F. Prytherch, Director of the U. S. Fish and Wildlife Service Laboratory, for facilities at Beaufort.

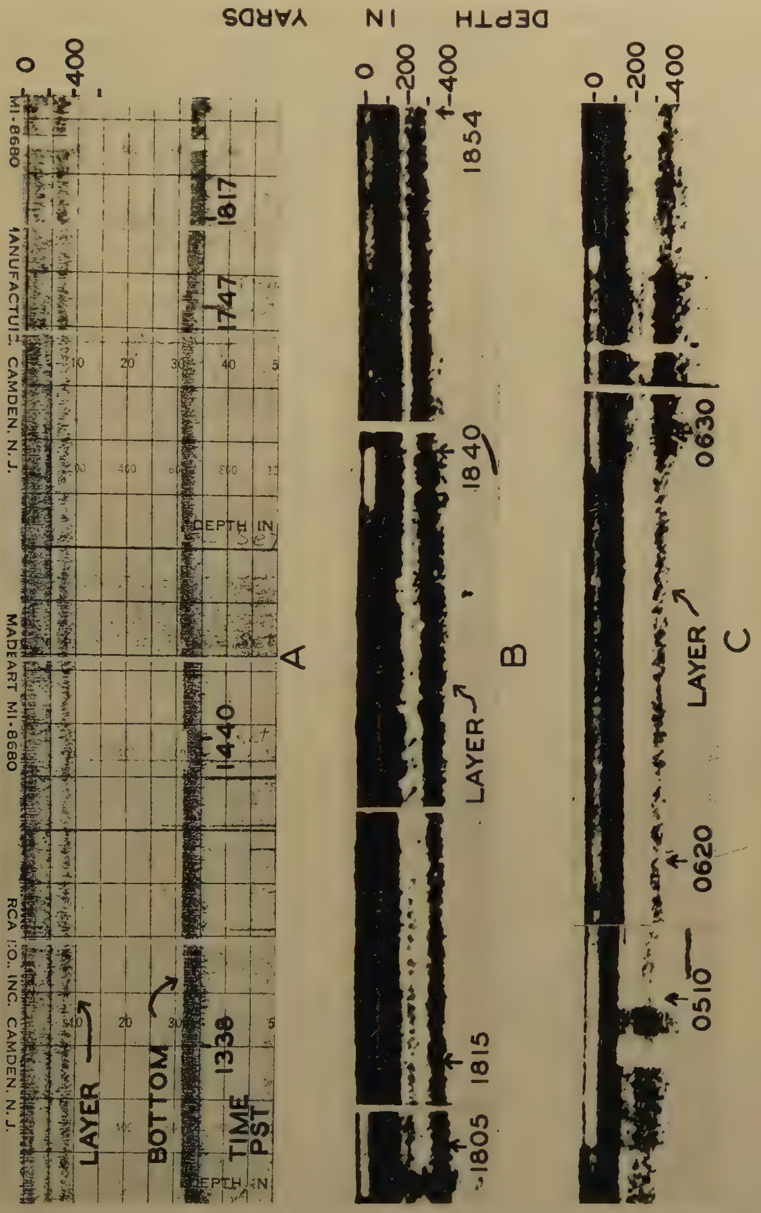


Figure 3. Fathograms of Deep Scattering Layer, 18 kc, E. W. Scripps, 1945, off San Diego: *A*, Layer during afternoon of June 26; *B*, Layer ascending August 18; *C*, Layer descending August 19.

This continues unabated for two to three hours and finally tapers off to individual outbursts at rare intervals. From tests and recordings of isolated fish in aquaria, it seems clear that the chorus consists of one or more of the following species of Sciaenidae: black croaker, *Sciaena saturna;* spotfin croaker, *Roncadro stearnsi;* yellowfin croaker, *Umbrina roncador.* Each of these fish produced vigorous froggy croaks which were nearly identical to those heard in the "chorus." In addition to croaking, *Sciaena saturna* produced a series of rapid clicking sounds. In the tests, none produced sounds similar to the soft background drumming.

Although tests of the black croaker demonstrated that individual fish are capable of producing more than one kind of sound, the soft drumming in the chorus may be produced by a species not tested, for it was never a part of the chorus during the first weeks in spring.

Further evidence which strengthens this view was obtained in July 1945 when listening alternately over stationary hydrophones, one placed at the pier and one at a distance of 2,000 feet from the pier. The sounds picked up on the distant hydrophone were dominated by the soft drumming, whereas on the nearer hydrophone the harsh croaks were dominant, with only a rare outburst of drumming. Thus it appears that two separate schools of different fish were involved.

The above is the only locality where a "croaker chorus" has been observed along the west coast, though individual outbursts are often heard elsewhere. Evening listening tests have been made at a few other locations with negative results, but it is highly probable that many other such choruses do hold forth at other localities south of Point Conception. The significance of this croaking in unison is not known, but it is difficult to believe that so marked and regular a phenomenon does not have some ecological significance.

REFLECTED SOUNDS

The Deep Scattering Layer

The "deep scattering layer" refers to a sound scattering layer within the sea. This layer, which was first observed in 1942 (C. F. Eyring, R. J. Christensen, and R. W. Raitt, 1948), may have a thickness of about 50 fathoms and be spread horizontally over vast areas at depths of 150 to 450 fathoms during the daytime. Outgoing signals of 18 kc frequencies are more or less scattered by reflections from bodies of scatterers that make up the layer. This results in more or less pronounced "false bottoms" being indicated on the recording fathometer tapes.

Biological Hypothesis

Precisely what these sound-reflecting bodies are is not known, but while working from the author's hypothesis that the deep scattering layer may be caused by stratification of either planktonic life or nektonic forms associated with the plankton, it was possible, in a 24-hour series of observations on June 26 and 27, 1945 aboard the E. W. Scripps, to show for the first time that the layer is not stationary at any depth but that it behaves biologically in performing vertical diurnal migrations in a manner paralleling the well known behavior of many planktonic animals (Anonymous, 1946).

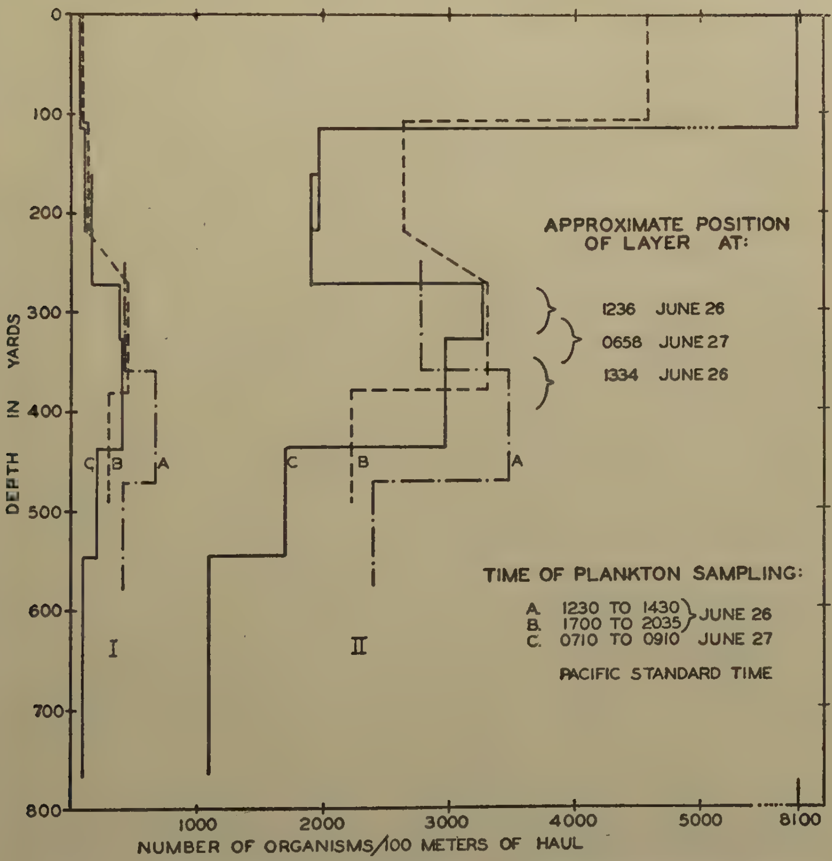


Figure 4. Vertical distribution of zooplankton. I. Animals greater than 2 mm. II. Animals 2 mm or less.

The observations were made about 20 miles off Point Loma where the depth is 660 fathoms. Fathograms were made of the layer from about noon June 26 to noon June 27 (Fig. 3). During this time it moved towards the surface in the evening and descended again the following morning, apparently under the stimulus of changing light intensity. It reached its maximum depth in early forenoon, suggesting that while light was the stimulus initiating the downward movement, some other factor (which might be temperature or pressure) became dominant in limiting the depth of descent. The partial or complete disappearance of the layer at night may result from a tendency of the scatterers to become dispersed as a result of the absence of the stimulus (light) which in the daytime causes them to accumulate in a relatively narrow zone.

Using the fathometer as a guide to the position of the scattering layer, a series of 14 vertical plankton hauls were made below, through and above the main layer (a faint secondary layer was indicated above the main layer). Nansen closing-type nets with No. OX and No. 8 bolting silk and a 40-cm mouth were used.[4]

In Fig. 4 is given the number of animals caught, adjusted to 100 meters of haul and shown as two size groups, *i. e.*, those greater than 2 mm in length and those 2 mm or less.

The daytime vertical distribution of the zooplankton was similar in general for the two days. The preponderance of smaller forms (mainly copepods 1 to 2 mm in length) in the 0 to 110 yard stratum was very striking. Below 110 yards the number of this size fell off markedly but increased again to a secondary maximum which in the early afternoon of June 26 was at about 360 to 475 yards, partly within and partly below the scattering layer. In late afternoon the maximum was somewhat higher and within the layer. On June 27, the secondary maximum of smaller forms was within and somewhat above the layer.

Animals of size greater than 2 mm in length showed a small increase within the scattering layer on both days. Most abundant of these were copepods ranging from 2 to 8 mm and including mainly *Calanus* and *Metridia*, but also *Heterorhabdus*, *Euchaeta*, *Eucalanus*, and *Rhincalanus*. The largest organisms caught were sagittae (total 35) ranging from 6 to 27 mm. Included also were a few euphausiids, ctenophores, medusae, polychaetes, and amphipods. Previous work in this area has often shown daytime swarms of euphausiids at depths below 200 yards. They are relatively strong swimmers and are capable of avoiding small nets during the daytime.

---

[4] It is obvious that in order to obtain larger samples and to minimize the effect of patchiness, horizontal tows should be resorted to in sampling the various strata. But samples collected in this manner at the deeper levels have thus far appeared unreliable because of the difficulty of maintaining the net within the narrow horizontal limits required.

The largest body cross section for animals other than the medusae was only 5 mm, and none of the animals was found to have air bubbles which would make them strong scatterers (cf. Raitt, 1948). Yet from these preliminary observations there appears to be some direct correlation of the planktonic animals with the scattering layer. If the size and acoustic properties of these forms are not suitable to scatter sound, the alternate explanation must then be that larger forms not sampled by the small, slowly moving net are responsible and that they occur more or less concurrently with the small planktonts upon which many of them presumably subsist. The larger fast swimming forms must include euphausiids, prawns, deep-sea fish, and squids, for it is well known that all these do frequent the waters within the range of the scattering layer.

The fact that correlation appears to be best with the larger planktonts may be of special significance. It suggests that many of the larger scatterers are not strictly filter-feeders (that is, do not filter out the small forms less than 2 mm in length with fine gill rakers, etc.) but probably select more or less individually the larger planktonts. The feeding habits of these deep-sea forms have received little attention, but the manner of feeding can be inferred largely from study of feeding structures. Fish such as the cyclothones, myctophids, and the hatchet fish (*Argyropelecus*) are not strictly of the filter-feeding types, nor are the squids. Hersey and Moore (1948) have recently shown the deep scattering layer to be widespread in the Atlantic Ocean, and present evidence that the euphausiids are perhaps among the planktonic forms involved as scatterers.

The rate at which the layer ascends or descends may be a measure of swimming speed and therefore may be an aid in deducing what animal groups might be involved in its formation. This rate apparently varies somewhat. In the June 26–27 observations the maximum rate of ascent appeared to be around six feet per minute. The record for August 18–19, 1945 aboard the E. W. SCRIPPS in the same area (Fig. 4 B and C) gives approximately the same rate of movement. According to the clearest records obtained by Dietz (1948) in an extensive survey of the geographic range of the layer, the rate of movement is 10 to 18 feet per minute. This rate of movement would exclude many slow swimming planktonic forms. However, the adult stages of such planktonic animals as euphausiids, chaetognaths, and amphipods can easily swim at this speed, and when under a directive stimulus, such as decreasing or increasing light, they might be expected to migrate vertically at a speed paralleling that observed in movements of the layer. Preliminary tests show that, for short distances at least, euphausiids are capable of swimming at rates ranging upward to 39 feet per minute, and the copepod *Eucalanus* as high as 30 feet per minute. The deep-sea fishes, such as the myctophids, sometimes caught in deep plankton hauls in this area, do not possess air bladders which would limit the speed of their vertical movements because of change in the hydrostatic pressure. At night they are also commonly caught at the immediate surface.

The vast extent of the layer (sometimes present for a whole day's run of the ship) presupposes an abundant and rather uniformly scattered community that shows up on the fathogram as being more widespread than schools of fish or squid commonly observed near the surface. Squid have been suggested as the probable cause of scattering (Lyman, 1948), but while they may be expected to be a part of the scattering layer along with other animals, the type of squid known to be abundant, i. e., *Loligo*, has been observed to school at the surface. If this habit is not abandoned when they are present at great depths, then it might be expected that they would appear as limited patches on the fathograms. For example, Tester (1943), in a study of reflected sounds applied to fisheries, has shown how schools of herring in shallow water can be detected as patches on the fathograms.

The acoustical evidence does not readily lead to the assumption that only one kind of animal is involved. On the other hand, Russell (1928) has reported a different migratory response for different age groups and sexes of copepods.

The daytime formation of two scattering layers, or the splitting of the layer, with one part migrating upward with increased darkness while the other remains stationary, as shown by Dietz (1948), presents complicated and intriguing problems for the ecologist. Do the layers represent segregations of different species, stages of development, or sexes?

These are questions that remain to be answered, but with the aid of acoustic gear as a guide to collecting the solution should be greatly simplified.

SUMMARY

Underwater sounds produced by animals or reflected by them can be used in studying concentrations and behavior of certain marine populations.

Biological Sounds. Crustacea and fishes are the main sound producers. Crustacean sounds consist mainly of high frequency crackle produced by snapping shrimp. With directional sound equipment, their habitats can be detected to distances of 2,000 yards or more.

Studies of the continuity and sound spectrum level of shrimp may give evidence pertaining to stability of populations, their diurnal activities, geographical distribution, and type of bottom and depth of water preferred.

Observations on fish noises have given less useful ecological data. They do, however, demonstrate a regular and pronounced seasonal and diurnal habit of certain Sciaenidae to form localized choruses, as contrasted with isolated solos by the toadfish (*Opsanus tau*).

Reflected Sounds. The "deep scattering layer" has been shown by its diurnal vertical migrations to be biological in nature. The partial reflection of fathometer signals in this layer promises to be a useful ecological tool in studying the organisms involved.

While using sound as a guide, a preliminary study in 1945 showed a positive correlation with plankton stratification and the depth of the scattering layer. This relationship may be partly secondary, since the more effective scatterers may be larger forms that subsist upon and migrate more or less concurrently with the plankton.

REFERENCES

ANONYMOUS

1946. Stratification of sound scatterers in the ocean. Sonar Data Div., Univ. Calif., Div. War Res., Rep. No. M397.

BURKENROAD, M. D.

1931. Notes on the sound producing marine fishes of Louisiana. Copeia, No. 1: 20–28.

DIETZ, R. S.

1948. Deep scattering layer in the Pacific and Antarctic Oceans. J. Mar. Res., *7* (3): 430-442.

DOBRIN, M. B.

1947. Measurements of underwater noise produced by marine life. Science, *105* (2714): 19–23.

EYRING, C. F, R. J. CHRISTENSEN AND R. W. RAITT

1948. Reverberation in the sea. J. acoust. Soc. Amer., 20: 462–475.

HAY, W. P. AND C. A. SHORE

1915–16. The decapod crustaceans of Beaufort, N. C., and surrounding region. Bull. U. S. Bur. Fish., 35: 371–475.

HERSEY, J. B. AND H. B. MOORE

1948. Progress report on scattering layer observations in the Atlantic Ocean. Trans. Amer. geophys. Un., *29:* 341–354.

HILDEBRAND, S. F. AND LOUELLA E. CABLE

1930. Development and life history of fourteen teleostean fishes at Beaufort, N. C. Bull. U. S. Bur. Fish., 46: 383–488.

Johnson, M. W.

1943a. Underwater sounds of biological origin. Univ. Calif., Div. War Res., Rep. No. U28. 20 pp.

1943b. Preliminary survey of certain biological underwater sounds on the east coast of North America. Univ. Calif., Div. War Res., Rep. No. U63.

Johnson, M. W., F. A. Everest and R. W. Young

1947. The role of snapping shrimp (*Crangon* and *Synalpheus*) in the production of underwater noise in the sea. Biol. Bull. Woods Hole, *93* (2): 122–138.

Knudsen, V. O., R. S. Alford and J. W. Emling

1948. Underwater ambient noise. J. Mar. Res., *7* (3): 410-429.

Loye, D. P. and D. A. Proudfoot

1946. Underwater noise due to marine life. J. acoust. Soc. Amer., *18* (2): 446–449.

Lyman, John

1948. The sea phantom bottom. Sci. Mon., N. Y., *66* (1): 87–88.

Radcliffe, Lewis

1914. The offshore fishing grounds of North Carolina. Econ. Circ. U. S. Bur. Fish., No. 8: 1–6.

Raitt, R. W.

1948. Sound scatterers in the sea. J. Mar. Res., *7* (3): 393-409.

Russell, F. S.

1928. The vertical distribution of macroplankton. VII. Observations on the behavior of *Calanus finmarchicus*. Mar. biol. Ass. U. K., N. S. *15:* 429–454.

Tester, A. L.

1943. Use of the echo-sounder to locate herring in British Columbia waters. Bull. Fish. Res. Bd. Canad., 63: 3–21.

CHANGES IN THE FISH FAUNA OF WESTERN NORTH AMERICA CORRELATED WITH CHANGES IN OCEAN TEMPERATURE[1]

By

CARL L. HUBBS

Scripps Institution or Oceanography
University of California
La Jolla, California

Many lines of evidence support the conclusion that temperature is a factor of prime importance in the distribution of marine fishes along the Pacific Coast of North America. The correlation between water temperature and the type of fish found is so definite that I feel justified in drawing inferences, from the known data of fish distribution, regarding ocean temperatures of the past and regarding the temperatures of regions for which the physical data are inadequate. Several patterns in the present distribution of fishes along the coast can be plausibly explained only on the theory that ocean temperatures fluctuated considerably, up and down, during recent geological time. Largely on the basis of warm-water forms reported from 1853 to 1860, I further infer that the ocean temperatures at that time, nearly one hundred years ago, were definitely higher in the southern half of California than they have been during any subsequent series of years.

Some call such two-way use of data unwarranted—a "vicious circle" —but I regard the method, when used with due caution and insight, as both proper and promising. Consistency in results and in explanations is a justification of the procedure. A similar approach was used in a recent correlation of hydrographic history and fish distribution in the Great Basin (Hubbs and Miller, 1948).

Even though I almost completely restrict the treatment zoologically to fishes, geographically to western North America, and temporally to the present and near-present, I must sail rapidly through a sea of extensive, varied and detailed data. In particular, the correlation between present temperatures and the present distribution must, for want of space, be treated sketchily. Chief emphasis in this paper is given to fish faunal changes associated with short-term (annual) and long-term changes in surface temperature within historic time.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 379.

REVIEW OF ZOOGEOGRAPHICAL EVIDENCE

The general transition of the fish fauna from Arctic to Tropic waters along the Pacific Coast illustrates well the correlation between temperature and distribution. The relative sharpness of the change in temperature at Point Conception is reflected in the rather rapid transition there between the fish fauna of southern California, with a large representation of subtropical types, and that of central California, with many subarctic forms.

The correlation between temperature and distribution is especially emphasized where the latitudinal transition is locally reversed. Thus, the general southward increase in temperatures and the increasingly tropical aspect of the fauna is reversed in northwestern Baja California, from near the California border southward through at least 2.5° of latitude. Here the offshore temperatures during the critical spring and summer months may run 1° C. cooler than near San Diego (according to unpublished data assembled under the direction of George F. McEwen). More significantly, the inshore temperatures of this region are in general unusually low, very strikingly so in foci of upwelling (McEwen, 1916). The validity of the areas of upwelling charted by McEwen (his plate 9) is being confirmed by recent data. Thus on August 8, 1948, the temperature of the surf was only 15.1° C. at or near the focus of the area of upwelling just south of the border, only 13.6° at Santo Tomás Anchorage south of Ensenada, and generally low in the intervening areas, except in Todos Santos Bay, where partial enclosure and shallow water produce a warm area, with temperatures on August 8 as high as 22.6° (at a point only 13 nautical miles north of Santo Tomás Anchorage). Furthermore, due to the prevalence of upwelling and the general lack of eddies, the temperatures off most of the coast of northwestern Baja California (Todos Santos and Sebastian Viscaino Bays excepted) can not be expected to attain the occasional heights experienced in southern California (as Harald U. Sverdrup pointed out in discussing this problem). In striking agreement with the oceanographic conditions, we are finding south of the international border, in the areas of upwelling, a very considerable representation of central California fish types. Some of the species, for instance the reef seaperch, *Amphigonopterus aurora*, appear to have a discontinuous range, broken in southern California and probably also in such warm areas of Baja California as Todos Santos Bay. Other fishes and some invertebrates (as the red abalone, *Haliotis rufescens*, and the large sea urchin *Strongylocentrotus franciscanus*) that occur near San Diego only in the cooler infratidal waters reappear between tides in the areas of upwelling in Baja California. Other species occur in shallower water in those areas. A strikingly parallel picture of algal distribution in northwestern Baja California has been discovered by E. Yale Dawson (personal communication). Our findings have been given brief advance notice, as by Hubbs (1946: 84), Barnhart and Hubbs (1946: 371) and Stephenson and Stephenson (1948: 7).

The prevalence along the Oregon and Washington coasts of upwelling, which lowers the inshore surface temperatures, seems to be associated with the absence there of certain fishes, such as the Californian reef klipfishes *Gibbonsia montereyensis* and *G. metzi*, that reappear in British Columbia. Other southern fishes that occur in British Columbia are listed by Clemens and Wilby (1946: 4). The extremely cold water just south of Monterey Bay is correlated with the occurrence there of such northern reef fishes as *Scytalina cerdale, Phytichthys chirus, Artedius harringtoni* and *Clinocottus embryum.*

Warmer bays have a far more southern fauna than the adjacent open coast. Thus we note the occurrence of a southern pipefish, *Syngnathus arcta*, and the abundance of the least seaperch, *Micrometrus minimus*, and of a southern flatfish, *Hypsopsetta guttulata*, in Elkhorn Slough and Tomales Bay, central California. Morro Bay has a particularly southern fauna, in harmony with the warmth of its waters, caused by the extensive shoals of blackish mud. Here occur the California killifish, *Fundulus parvipinnis*, otherwise seemingly absent north of Santa Barbara, the southern California subspecies of the staghorn sculpin, *Leptocottus armatus*, and the northernmost population of the pipefish *Syngnathus griseo-lineata leptorhyncha*, separated from the main range of this subspecies by another subspecies, *S. g. barbarae.*

The along-shore cooling effects of upwelling in Oregon and Washington are strikingly evident in the southward extension of the range of many subarctic fishes. Farther offshore, beyond the effects of upwelling, occur fishes of more southern types, such as the mola (*Mola mola*), albacore (*Germo alalunga*), bluefin tuna (*Thunnus thynnus*), and more southern types of sharks.

In the cases of more or less isothermal bathymetric distribution, the effect of temperature transcends that of light and a great variety of other factors. Thus certain forms that inhabit intertidal reefs in central California occur in the cooler infratidal waters in southern California. Examples are a klipfish, *Gibbonsia montereyensis*, and a rockfish, *Sebastodes chrysomelas*. Like the two invertebrates mentioned previously (*Haliotis rufescens* and *Strongylocentrotus franciscanus*), these two species reappear on certain intertidal reefs of Baja California where the water is unusually cool. A midshipman, *Porichthys notatus*, breeds in intertidal waters in central California but not in southern California. In general, fishes of wide latitudinal distribution occur in deeper water to the southward. Thus the ratfish, *Hydrolagus colliei*, lives close to shore in Puget Sound, but chiefly in about 50 to 100 fathoms near San Diego. The seven-gill cowshark (*Heptranchias maculatus*) is common in the shallow parts of San Francisco Bay, but off San Diego it is taken chiefly off the open coast at depths of about 50 fathoms. A rockfish, *Sebastodes mystinus*, abounds in central California close to the rocky shores and near the surface offshore, but is confined in southern California to rather deep water. It is hoped that supporting data will be marshalled in future publications.

The seasonal movements of fishes along the coast are clearly controlled by water temperatures. Thus marlin, *Tetrapturus mitsukurii*, do not appear off San Diego until or unless the surface temperatures approximate 70° F. Albacore, *Germo alalunga*, are looked for at temperatures approximating 60 to 65° F. The wide movements of the sardine, *Sardinops caerulea*, are also definitely seasonal.

Though the data have been presented so sketchily and though much of the detailed information remains to be gathered and collated, it seems safe to assert that the distribution of marine fishes along the Pacific Coast of North America is determined to a large degree by the sea temperatures. This assertion, of course, is bolstered by the close relations between temperature and distribution established for fishes in other parts of the world, and for other groups of organisms.

INFERENCES FROM FAUNAL EVIDENCE REGARDING PREHISTORIC CHANGES IN SEA TEMPERATURE

On the basis of present distribution it may be inferred that isotherms and fish faunas have been displaced both northward and southward during recent geological time. Unhappily the Pleistocene fish faunas are much too poorly known to bring to bear on this problem, in the way that warm-water and cool-water Pleistocene mollusk faunas have been cited to indicate such thermal and faunal shifts (Smith, 1919; Schenck and Keen, 1937; etc.).

The evidence for past northward shifts in ocean isotherms and fish distribution is found chiefly in the occurrence of southern forms northward, in appropriate waters beyond discontinuities in distribution. Reference has been made already to the occurrence in British Columbia of certain Californian fishes that do not live in the intervening cool regions of maximum upwelling; also, to the northern outliers of southern types in certain warmish bays, particularly Morro Bay. Since none of the southern relicts shows any evidence of differentiation, the past connections, probably during periods of warm climate, were presumably in late Quaternary time. The warm Postglacial period, generally recognized on the basis of pollen profiles and other evidence, suggests itself as the time when isotherms and fish faunas were displaced northward along the coast.

A small body of fish evidence and a larger amount of mollusk data from Indian kitchen middens support the view that displacements both northward and southward have taken place during Postglacial times.

The occurrence along the northwestern coast of Baja California of northern coastal relicts, likewise undifferentiated, may be taken to indicate a very late Pleistocene or early Recent southward displacement of isotherms and fish faunas. The presence of a rainbow trout, *Salmo gairdnerii nelsoni*, in certain streams of the region has been thought of as confirmatory evidence of such a southward shift, but the inference is not secure (Hubbs, 1946: 85), not nearly so convincing as the inference for past displacements that can be drawn from the presence in Formosa (Jordan and Oshima, 1919: 122; Jordan and McGregor, 1925: 137) of a land-locked salmon (*Oncorhynchus formosanus*).

A more ancient and more extensive southward displacement of isotherms and faunas is indicated by the occurrence of a rather large number of northern relicts in the upper part of the Gulf of California, where the summer water temperatures are high and the winter temperatures low. The types with northern affinities coexist with a somewhat depauperate Panamaic fish fauna. Most of the northern Gulf relicts and their respective Californian cognates do not now range into the almost strictly tropical waters about Cape San Lucas. The connection presumably existed at some ancient time, or more probably at several times. Since, according to published and unpublished studies, the level of differentiation grades from full-specific to racial or none—usually weakly specific or subspecific—and since southward displacements are generally accepted as having occurred during the Glacial periods, it seems logical to date the connection or connections as Pleistocene. The occurrence of a specifically differentiated trout of the *Salmo gairdnerii* series in the upper part of the Gila River system seems attributable to a similar displacement. Other northern fish relicts in various streams of the Pacific Coast drainage are probably also of Pleistocene derivation (data in part cited by Hubbs and Miller, 1948).

LONG-PERIOD NORTHWARD DISPLACEMENTS OF FISH FAUNAS WITHIN HISTORIC TIME

In agreement with other lines of evidence, data on changes in the distribution of fishes along the Pacific Coast suggest that climatic oscillations have persisted into historic time. There are strong indications that the ocean temperatures from 1850 (or earlier) to about 1870 were definitely warmer in the southern half of California than during subsequent decades, and there are weaker suggestions of a reverse change in temperature and fauna during the past two or three decades.

The fish fauna of San Diego, as sampled from 1853 to 1860, particularly by the Pacific Railroad Survey from 1853 to 1857 (Girard, 1858; etc.), was definitely more tropical than that of any subsequent decade. Of the 30 odd species reported, six (about 20%) do not now occur so far north or have been so rare recently that one certainly would not expect any to be caught at present by such incomplete and superficial collecting as that of the 1850's and 1860's. Possible sources of gross error have been disregarded after due consideration. The Pacific Railroad Survey secured at San Diego: five specimens of the giant seahorse, *Hippocampus ingens*, few of which were collected north of Magdalena Bay from then until 1880 and none since that year; four examples of a presumably tropical pipefish, *Dermatostethus punctipinne*, which has never been secured again, anywhere; several individuals of an unidentified triggerfish (of the tropical family Balistidae), which seemingly has not been collected again in the eastern Pacific; two spadefish, *Chaetodipterus zonatus*, of which the only subsequent California record is for a specimen taken in the extremely warm summer of 1931 (p. 467); a puffer, *Sphoeroides politus*, of which only one doubtful example has been secured in California since 1880 and few before that year; and a jack, *Caranx caballus*, of which single specimens have been reported from California in 1857, 1924, and 1945. The only northern species reported from San Diego by Girard was the seaperch *Taeniotoca lateralis*, but this species was represented by a single adult; this species ranges commonly south to the Santa Barbara Islands and it has been collected recently near San Pedro and in northern Baja California.

Nearly as impressive is the evidence that the fish fauna at Monterey during the first two decades of the second half of the nineteenth century was made up to a considerable degree of fishes now characteristic of the relatively warm waters south of Point Conception. Only about twenty species were reported from Monterey by the Pacific Railroad Survey, but the list includes five or six kinds that do not now occur so far north or are so rare as not to be collectable by any such dribbling survey as that of the 1850's. These southern types are the sand bass, *Paralabrax nebulifer*, the California sculpin, *Scorpaena guttata*, the bay blenny, *Hypsoblennius gentilis*, the garibaldi, *Hypsypops rubicunda*, the horned shark, *Heterodontus francisci*, and perhaps the guitarfish, *Rhinobatos productus* (recorded as from either Monterey or San Francisco). The evidence that the California flyingfish, *Cypselurus californicus*, occurred prior to 1864 north as far as Santa Cruz on Monterey Bay has been cited (Hubbs and Kampa, 1946: 209) as an example of former occurrence of southern forms north of their present range. As noted by Steinbeck and Ricketts (1941: 48) and by Schmitt (1921: 163), *Pleuroncodes planiceps*, a very conspicuous crustacean that swarms at the surface in warm water, was washed ashore in considerable numbers in March 1859 at Monterey, hundreds of miles north of its usual present range. Ordinarily this species is uncommon north of central Baja California, though it penetrated into southern California in abundance in the spring and summer of 1941 (notes in California State Fisheries Laboratory).

For the region of San Francisco, in contrast, the evidence (more extensive than for the other ports) indicates a relatively stable fauna throughout the last hundred years. Approximately one hundred species have been reported as collected about San Francisco from 1853 to about 1860 (here much local collecting, particularly by W. O. Ayres, supplemented that of the Railroad Survey). With about seven exceptions, the species reported are still rather common about San Francisco. One of the seven, the opaleye, *Girella nigricans*, has not since been collected about San Francisco, but adults straggle to Monterey, and Ayres had only one example from San Francisco. Four (namely *Squatina californica*, *Rhinobatos productus*, *Synodus lucioceps* and *Stereolepis gigas*) are now very seldom found north of Point Conception, which further suggests a slight southward facies, though all of the species were reported as rare in the 1853–1860 period. Partly counterbalancing this indication of a southern facies were the early records of two northern species for San Francisco. Girard and Ayres each reported from there a snakeblenny, *Lumpenus gracilis*, which has never since been collected south of Puget Sound, and Ayres had one, the sanddigger, *Trichodon trichodon*, which has since been taken only once at San Francisco.

The relatively large number of subtropical types that were taken about San Diego and of southern California types that were secured about Monterey, from 1853 to about 1860, and the stationary habits of these species, make it appear very improbable that the northward occurrences represented merely single-season waves, such as are mentioned below. A further reason for believing that the warm-water faunas were long-lasting and real is that a considerable representation of southern types seems to have persisted through the 1870's until 1880, when Jordan and Gilbert conducted the most thorough reconnaissance of the century of the West Coast fishes. A study of the reports by Jordan and Gilbert (1881 a, b) and by their contemporaries indicates that the fauna from San Diego to San Francisco had a slightly more southern tinge than at present, despite the fact that 1880 was an unusually cold year (Figs. 1–3, 5). No particularly northern type was reported for 1880, except possibly a cod, *Theragra chalcogramma* (recorded as rarely taken about San Francisco), but in contrast a considerable number of distinctively southern types were reported for that year. A Spanish mackerel, *Scomberomorus concolor*, occurred in some numbers in Monterey Bay then and during the following several years, but it is now excessively rare even in southern California. *Cynoscion parvipinnis* was common in southern California, where it now seems to have been absent for many years, though it remains abundant in central Baja California. The barracuda, *Sphyraena argentea*, was said to be especially abundant in Monterey Bay, whereas now it lives chiefly in southern California and southward.

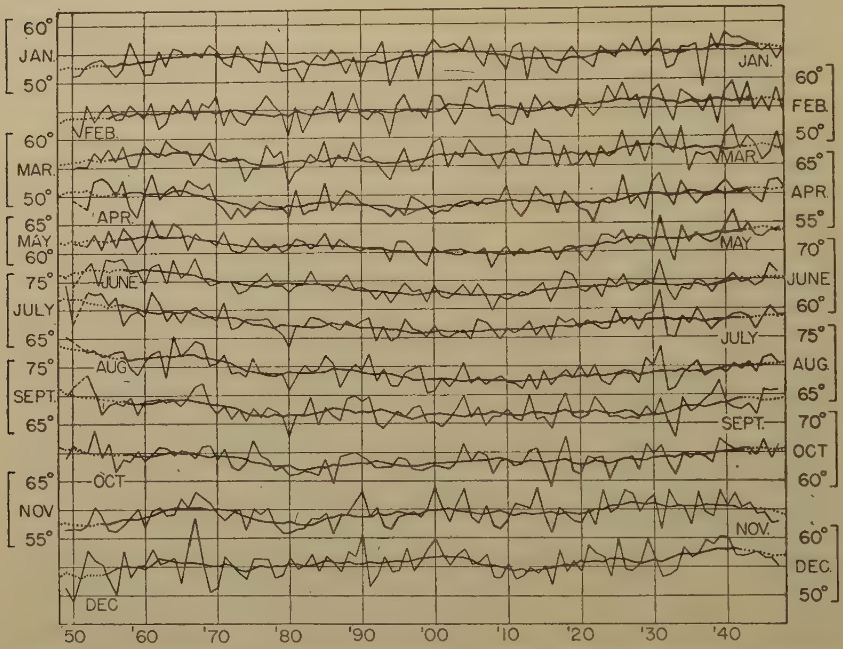


Figure 1. Fluctuations of mean air temperatures at San Diego from 1849–50 to 1947–48, for each month, with trends computed by running averages of 11. Data from World Weather Records (Smithson. misc. Coll., *79*, 1927: 934–935; *90*, 1934: 347; *105*, 1947: 373) and from Climatological Data (U. S. Weath. Bur., vols. 45–52, 1941–1948); all data corrected to means of 24 hours.

The bonito, *Sarda lineolata*, now chiefly southern, was reported to be very abundant everywhere in summer. The elasmobranch fauna as reported seems distinctly more southern than at present.

Several tropical types reported as occurring rarely in southern California in the last third of the nineteenth century possibly also represent holdovers from the warm-water fauna of the middle of the last century. Among these are three that have not since been taken in the state: *Eucinostomus argenteus*, *Chilomycterus affinis* and *Diodon hystrix*.

During the last two decades a considerable number of tropical types have been collected in southern California, suggesting the possibility that a slightly more tropical assemblage is becoming re-established. However, the captures are attributable, in part at least, to more thorough collecting, to chance migrations, and to the single-season northward dispersals that are discussed in the following section. Four bottom fishes previously unknown north of central Baja California, however, have been found recently in some numbers in San Diego County, with indications of recent establishment. These are the zebraperch, *Hermosilla azurea*, a porgy, *Calamus brachysomus*, and two giant groupers, *Mycteroperca xenarcha* and *M. jordani*.

SINGLE-SEASON NORTHWARD DISPERSALS OF FISH FAUNA

During occasional warm years, unusual numbers of free-swimming fishes and other pelagic animals appear along the Pacific Coast far north of their customary range. A year particularly noted for this phenomenon was 1926 (Hubbs and Schultz, 1929). Southern types, including the siphonophore *Velella*, abounded as far north as British Columbia. *Mola mola* appeared off the British Columbia and Washington coasts. Anchovies, *Engraulis m. mordax*, were unexpectedly abundant in the mouth of the Columbia River. Albacore, *Germo alalunga*, were caught near the Oregon shore for the first time on record and the catch off central California increased sharply. Mackerel, *Pneumatophorus japonicus diego*, and jacksmelt, *Atherinopsis californiensis*, previously unrecognized there, appeared in numbers in Coos Bay. A thresher, *Alopias vulpinus*, was also taken there, as a novelty, and many hammerheads (*Sphyrna* sp.) were caught in southern California.

Five years later, in 1931, an unusual number of southern fish strays were captured along the coast of southern California (Walford, 1931). In 1941 red crabs, *Pleuroncodes planiceps*, abounded off Los Coronados and Santa Catalina islands, far north of their usual range. In British

Columbia there was an intrusion of southern forms, as also in 1936 (John L. Hart, personal communication).

Similar reports stem from off the Oregon, Washington and British Columbia coasts for the summer of 1947, when July was exceptionally clear and warm. According to fishing reports, swordfish and even flyingfish appeared for the first time off Oregon, along with an especial abundance of albacore and numbers of bluefin tuna. A "yellowtail," presumably *Seriola dorsalis*, was caught off southern Washington, far north of the previously known northern limit of range. California pompano, *Palometa simillima*, appeared in unprecedented abundance in British Columbia (Hart, in press).

The appearance of southern fish on the Pacific Coast far to the north of their usual range, during warm years, has usually been attributed to the northward dispersal of individuals. Movement of this sort may be involved, at least for such fish as the sardine, *Sardinops caerulea*, and the mackerel, *Pneumatophorus japonicus diego*, that are known to migrate regularly up and down the coast. An alternative explanation, for animals with a very short life cycle, is that local populations in the north, too small to be conspicuous, may suddenly erupt when temperatures rise. More commonly, however, it is probable that the warm-water types merely move closer to shore from the warmer waters that lie outside the coastal zone of upwelling. Thus the albacore that occasionally appear close to the Oregon and Washington coast probably come chiefly from the population of the offshore, warm "blue water," rather than from southern coastal waters. This is even more likely true of *Mola* and *Velella*. The dispersal may be essentially short and transverse, like the seasonal migration of many western birds which, instead of undertaking long north and south movements, merely fly back and forth between mountains and valleys.

Whatever the course of movement may be, there is little doubt that in years when the surface temperatures are high, particularly during the spring and summer months, southern types of animals occur near the Pacific Coast far north of their usual limit of range. However, there is no regular northward dispersal comparable to that effected along the Atlantic Coast by the Gulf Stream.

TEMPERATURE DATA BEARING ON CHANGES IN FISH FAUNA

Water temperatures are available for correlation with some of these short-period shifts in fish fauna along the Pacific Coast but are inadequate to explain unequivocally any of the long-term changes. Air-temperature data, within limits, can be used as a substitute for lacking or inadequate water-temperature data, or can be used to corroborate limited data on sea temperatures, not only for current and recent years but also for the latter half of the nineteenth century. Estimates of prehistoric changes in the ocean temperatures rest on our knowledge of the temperatures of waters now inhabited by the same or similar faunal assemblages and will probably continue to rest on that basis, though some corroboration may be expected from intensive study of middens and fossil deposits and eventually from correlations of the sea-levels and associated deposits with Pleistocene history.

The noteworthy northward displacement of the marine fauna in 1926 was clearly due to exceptionally high sea temperatures along the whole length of temperate North America during the first half of that year. Hubbs and Schultz (1929) reached this conclusion from an analysis of sea temperatures at Pacific Grove, California, and of air temperatures at coastwise stations in Oregon and Washington. Sea

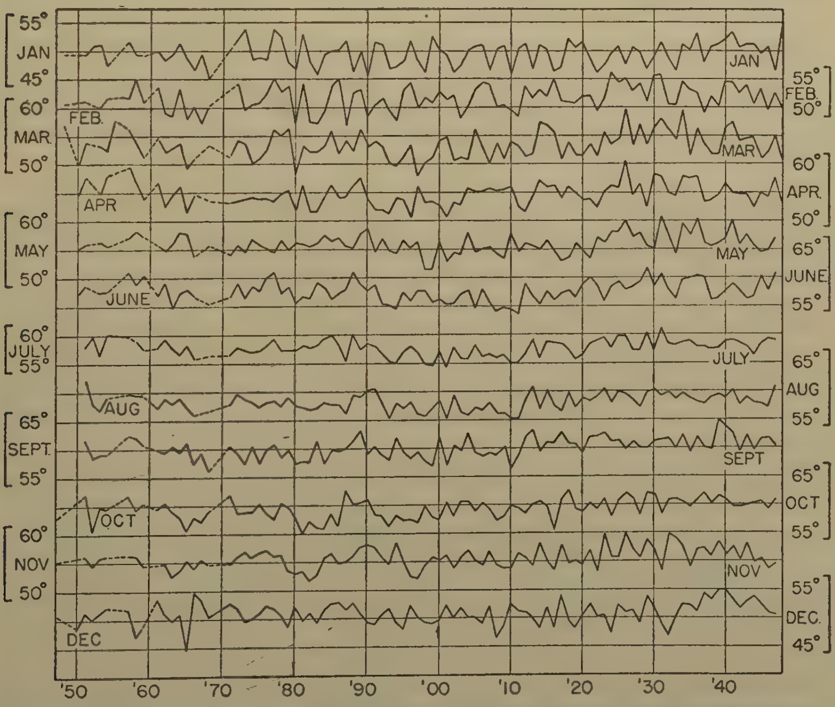


Figure 2. Fluctuations of mean air temperatures at San Francisco from 1847 (in part) to 1947–48, for each month. Data from World Weather Records (Smithson. misc. Coll., *79*, 1927: 939–940; *90*, 1934: 348; *105*, 1947: 374) and from Climatological Data (U. S. Weath. Bur., vols. 45–52, 1941–1948); all data corrected to means of 24 hours.

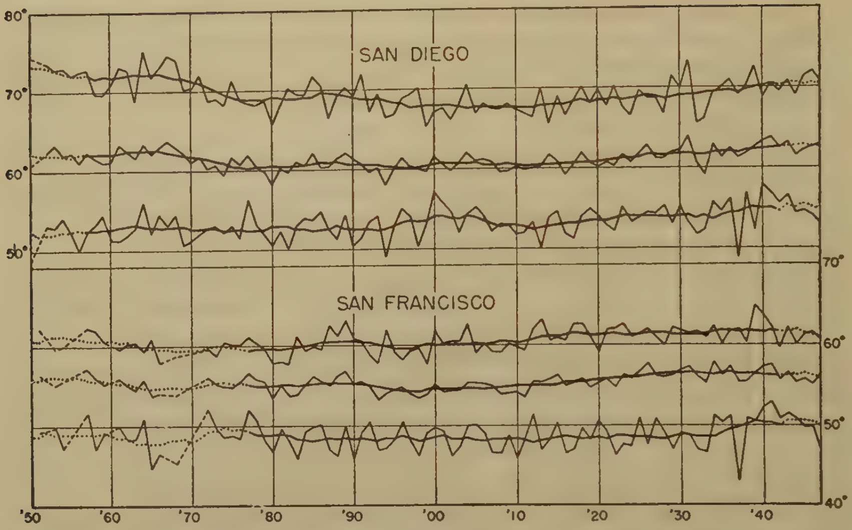


Figure 3. Fluctuations of mean temperatures at San Diego and at San Francisco, with trends computed by running averages of 11. Shown for each city are the means for the warmest month of each given year; the means for the year; the means for the coolest month. Source of data as for Figs. 1 and 2.

temperatures were high during the first half of 1926 at other stations in California (in general the correlation between stations is very close). Thus, at San Diego, the annual mean was unusually high in 1926 (Figs. 3, 5) as a result of extremely warm water during the early months of the year (Fig. 4). A similar pattern was displayed by the air temperatures at San Diego (Figs. 1, 3–5) and at San Francisco (Figs. 2, 3). At Victoria, British Columbia, the air temperatures were at or near record height from January through June and in November; about median in the other months.

During most of 1931, when many tropical fishes appeared off San Pedro, California, as well as during the latter part of 1930, the air temperatures were definitely above normal at Los Angeles and at San Diego (Walford, 1931). Sea temperatures at La Jolla (Fig. 4), closely corresponding with the air records for San Diego (Figs. 1, 3–5), were likewise extremely high from January through September 1931. The temperature increase in 1931 was less widespread than in 1926 and no record seems to have been made of any general appearance of southern forms in 1931 at any point in California north of the southern part of the state. At San Francisco the air temperatures ran about normal in most months (Figs. 2, 3). In British Columbia, however, air and sea temperatures were above normal and there was an intrusion of southern forms (John L. Hart, personal communication).

During the early months of 1941 the air temperatures were again very high (Fig. 1) at San Diego and the mean air temperature there,

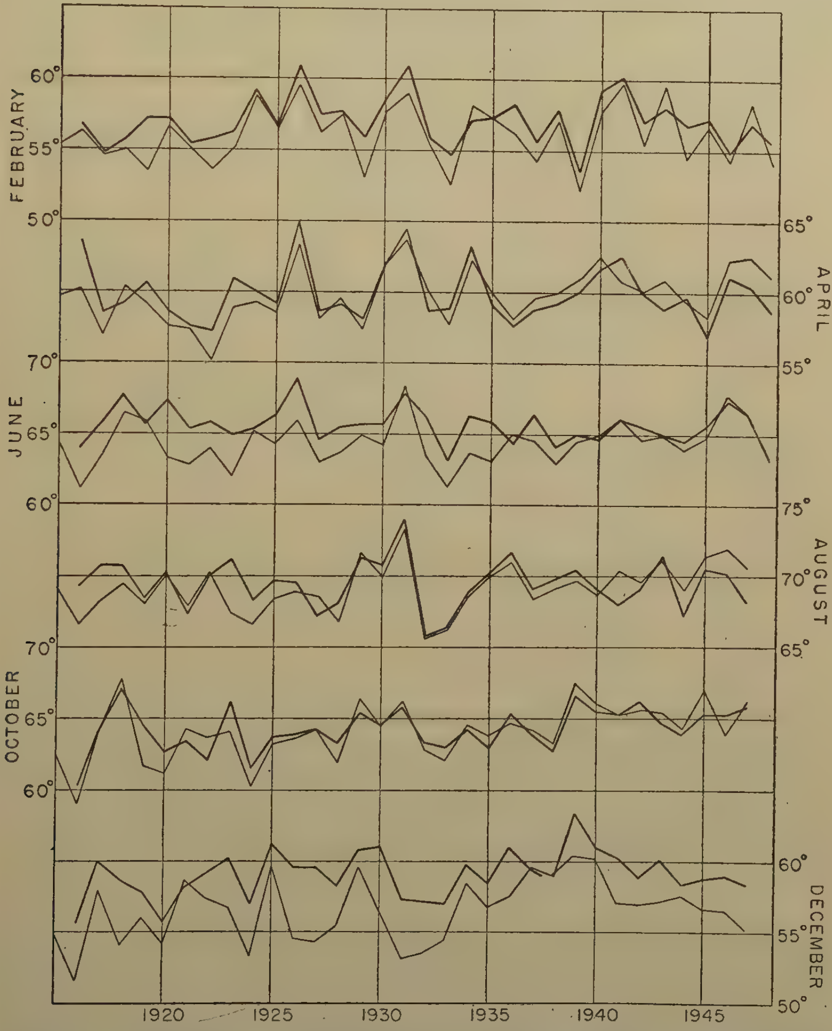


Figure 4. Fluctuations of mean temperatures for alternate months. Lighter lines represent air temperatures at San Diego from 1915 to 1947–48 (source of data same as for Figs. 1–3); darker lines represent the monthly means of the sea-surface temperatures taken at 7:00 A.M. each day at the end of Scripps Pier, La Jolla.

as well as the mean sea temperature at La Jolla, was very high (Fig. 5). This is the year when the warm-water red crab, *Pleuroncodes planiceps*, abounded during the spring and summer off southern California (p. 467).

Data collected by John L. Hart of the Fisheries Research Board of Canada, from the albacore fishing grounds of the northwest coast and from the Nootka, British Columbia, lighthouse, confirm fishermen's reports of unusually high summer temperatures in 1947, when southern forms were common in the north.

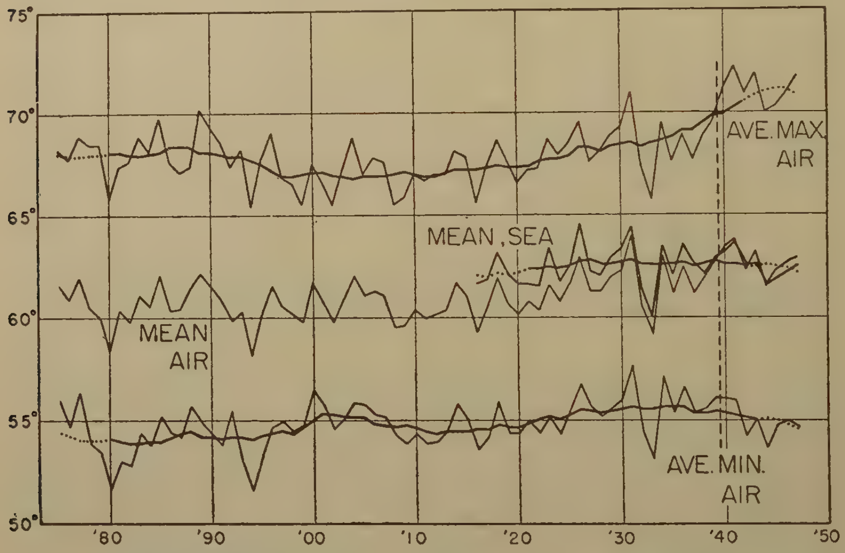


Figure 5. Comparison of fluctuations of minimum, mean and maximum air temperatures at San Diego from 1875 through 1947 with trend of sea temperatures at Scripps Pier, La Jolla, from 1916 through 1947. Annual means are plotted, with trends computed by running averages of 11 (the trend for the mean air temperature is omitted, but is graphed on Fig. 3). The upper and lower curves represent the annual averages of the daily maximum and the daily minimum air temperatures. The mean air temperatures are corrected to means of 24 hours. The vertical dashed line represents the time when the weather station was moved to a location with higher maximum and slightly lower minimum temperatures.

Greater complexities and uncertainties are involved in the effort to correlate long-term changes in the fish fauna with changes in the ocean temperatures. Air temperature data from the weather records (Figs. 1–5) stand in striking agreement with the limited though impressive data on fish faunal changes. Air and sea temperatures along the Pacific Coast are very closely correlated, as is generally known and confirmed by the data presented below, but doubts as yet unresolved pertain to the correlation of air temperature records of the last century with contemporary changes in sea temperatures.

The air-temperature data for San Diego (Figs. 1, 3–5) strongly suggest long-term climatic changes that correlate nicely with the observed faunal changes. The mean temperature seems to have been high during the early years of record, beginning in 1849–50 and continuing through the 1850 and most of the 1860 decades; that is, during the period when the fish faunas at San Diego and at Monterey were distinctly more southern than during any of the subsequent decades. From just before 1870 the general indicated trend was toward cooler weather until about 1910, after which there seems to have been a moderate reversion toward warmer conditions. In possible correlation, there is evidence—rather slight to be sure—that some subtropical fishes are becoming established in the San Diego area.

Considering the data by months (Fig. 1) we find that the changes in the trend for the mean annual temperature at San Diego characterize also the trend for each month from April through October. For two

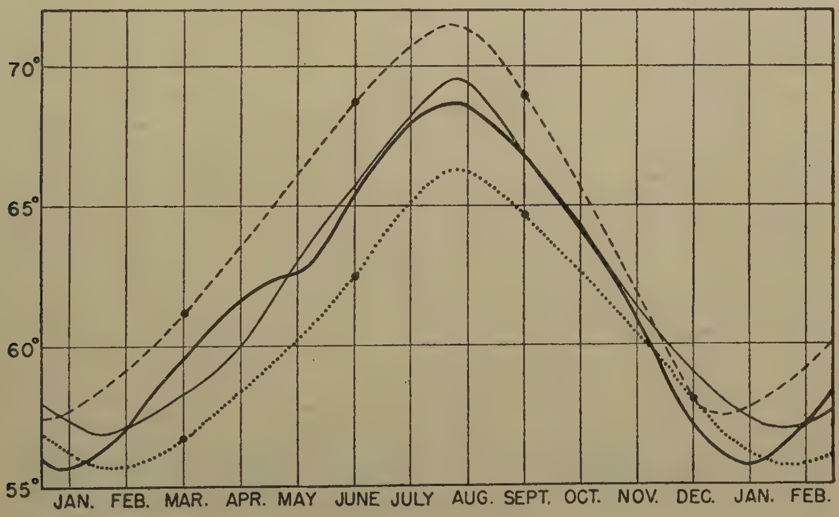


Figure 6. Variation through the year of the average sea-surface temperature about San Diego. The lighter solid line represents the trend at Scripps Pier, La Jolla, from 1916 through 1947. The heavier solid line shows the data for San Diego Bay entrance from 1854 to 1860 (probably depressed during late spring and summer for reasons given in the text). The dotted line indicates roughly the temperature cycle at present to the south of the bay entrance, computed from observed quarterly differences from the Scripps Pier data for the same year (1946–47). The dashed line is a maximum inference of the sea-surface temperatures at La Jolla, based on the assumption that the bay-entrance water came from the southward in 1854–1860, rather than from off Pt. Loma as it apparently does at present.

reasons this evidence is of particular importance in the present connection. In the first place, the months from April to October comprise the period of rapidly increasing and high sea temperatures in southern California (Fig. 6), that is, the season when the warmer-water forms of life reproduce and pass through critical stages of development. In the second place—a point that Martin W. Johnson has suggested—the indicated period of increased warmth was long enough to be of real biological significance. The fact that the months from November through March were not warmer from 1850 to 1870 than during subsequent decades, perhaps even cooler, might not have affected the northward displacement of the fauna nearly a century ago, because the region from which the southern forms would have been derived—southern Baja California—is characterized by cool winter sea temperatures as well as by warm summer seas; the subtropical forms occurring there must be adapted to cool winters.

At San Francisco (Figs. 1, 3), where there seems to have been only a slight change in the fish fauna over the last century, the air temperature data since about 1850 do not indicate a marked long-term trend downward from the 1850–1870 decades to about 1910. There seems to have been a slight downward trend, with notable irregularities, in the months from April through August.

That the air and sea temperatures along the Pacific Coast are correlated very closely is well known from the work of George F. McEwen and others and is strongly confirmed by recent analyses. There are some irregularities, for example, when offshore winds from the desert (Santa Anas) markedly raise the air temperatures but cause upwelling of cool water. Comparing the mean monthly air temperatures at San Diego with the mean monthly morning (7:00 A. M.) sea temperatures at the Scripps Institution Pier (Fig. 4, showing the data for alternate months), we find a very close agreement for the period from 1916 to 1947–48. Major discrepancies occur in certain months, notably January, June and December, but in other months the fluctuations in air and sea temperatures correspond remarkably well. For periods of several years the means and the changes from year to year may be almost identical. The curves for the annual means from 1916 through 1947 (Fig. 5) parallel one another with great precision. The mean sea temperature averaged 0.64° C. higher than the mean air temperature from 1916 through 1939 and 0.10° lower from 1940 through 1947, during which period the air temperatures were taken at a new, warmer location, as explained below.

The correlation coefficient (*c*) between the average air and sea temperatures is very high. Between the mean annual sea temperatures at La Jolla and the mean annual daily minimum air temperatures at San

Diego, for the period from 1916 through 1946, $c = 0.87$ (the minimum temperatures were only slightly reduced by the move of the Weather Bureau station). There was a similar high correlation between the mean sea temperatures and the annual average of the daily maximum temperatures: $c = 0.87$ for the period from 1916 through 1939 and $c = 0.86$ for the period from 1940 through 1946. For five years, from 1940 to 1944, the correlation coefficient between the monthly average daily sea temperatures at La Jolla and the monthly average of the daily evening radiosound surface air temperatures (approximately 2° F. higher than the minimum air temperature) at San Diego reached the astonishingly high value of 0.98. These data were computed by Dale F. Leipper of the Scripps Institution and were kindly made available by him.

Since the air and sea temperatures are so very closely correlated about San Diego, one might assume that the high air temperature records for the 1850 and 1860 decades and the low air temperature from about 1870 to about 1910, with a slight increase since, might be taken as a clear index of changing sea temperatures. Possible errors, however, are involved. From July 1849 through September 1871, the records were kept by the Army Medical Corps at various locations, by methods and with instruments that are not recorded in detail (Carpenter, 1913: Chap. I). The exact location of the station was changed several times, but fortunately the station remained at about the same elevation and within a one-half mile radius, from 1860 (about the middle of the indicated warm period) through January 1940 (Carpenter, 1913, and Dean Blake, personal communication). Thus the warm summers, the cool winters and the wide difference between the warmest and coolest months, indicated for the early years of the San Diego records (Fig. 1), can hardly be attributed to a change in the location of the station. It is suspicious that the temperature records were generally high until taken by the standard methods of the Weather Service (first under the Army Signal Service), beginning October 1, 1871. Bespeaking the significance of the indications of a prolonged warm period from 1849 through 1868, however, are the facts (1) that moderate temperatures were reported in 1869 and 1870, before the Weather Service took charge, and (2) that, with fluctuations, the trend continued downward after the methods were standardized. Considering also the striking agreement between the recorded weather data and the indicated change in fish fauna, I regard the balance of evidence as favoring the view that the trend of both air and sea temperature was downward from about 1870 to about 1910.

However, some actual data on water temperatures from 1854 to 1860 do not yield unequivocable support for the view that the sea temperatures of the southern California coast were higher then than at present (Fig. 6). These data were taken at noon almost every day from May 26, 1854, through October 31, 1860, by Andrew Cassidy, Collector of Customs and operator of the tide-gauge station that was established by the United States Pacific Railroad Survey. They are part of a set of meteorological observations made at the tide station in the entrance to San Diego Bay. The location was on the Point Loma shore north of Ballast Point, at the site long occupied by the Quarantine Station and now serving as the dock area of the Navy Electronics Laboratory (and of Scripps Institution of Oceanography). These observations, in a set of notebooks on file in the Junipero Serra Museum (of the San Diego Historical Society), were made available by Director John Davidson.

The annual mean of these temperatures taken at the San Diego Bay entrance from 1854 to 1860 are almost identical with the annual mean of the readings taken at 7:00 A. M. every morning from 1916 through 1947 at the Scripps Pier: 16.83° C. (62.30° F.) versus 16.98° C. The monthly means for 1854-60 fluctuate about the means for 1916–47: they were nearly 2° F. cooler in December and January (in correspondence with the cooler air records for these months during the same period) and nearly 1° F. cooler in August. In the probably critical spring months of March and April the means for the early years were 1.18° and 1.75° F. higher than for recent years. Though the data covered only six years, it seems probable that the winter sea temperatures of nearly a century ago were cooler, the early spring temperatures warmer, than in recent years.

It is assumed that the 1854–60 data represent ocean rather than bay conditions and that the water movement in the bay entrance was back and forth from some point near the harbor entrance. If bay conditions had ruled, the temperatures would have fluctuated widely with the tides, so that the noon readings would have varied with the progression of the tides. No such variation is apparent. In fact the readings are remarkably constant. The fluctuations within one month varied from 1° to 9° F., with an average of only 4.5° F. for the 77 months involved.

On their face value the 1854–60 data indicate, however, that the late spring, summer and fall sea temperatures were about the same as, or a little cooler than, at present, but there are reasons for thinking that the two sets of data are not comparable and that the ocean temperatures were warmer in 1854–60 throughout the spring and summer than they are now. To be sure, duplicate temperatures taken at the same spot in July and August, 1948, average about the same as the Scripps Pier records, in fact slightly higher, but the hydrographic conditions were greatly altered by the dredging of the bay and by the construction of the jetty on Zuñiga Shoal (between 1893 and 1904, according to Hertlein and Grant, 1944: 26). These actions resulted in the widening, deepening and straightening of the harbor entrance. In July and August 1948 the water surging back and forth through the entrance appeared to stem from the water near the tip of Point Loma where the along-shore water temperatures closely approximate those that extend north past the Scripps Pier (about 12 land miles due north of the harbor entrance) to the headlands near Laguna Beach. In September 1948 the bay-entrance temperatures averaged about 0.5° C. lower than those at La Jolla.

From just east and south of the harbor-entrance jetty definitely cooler temperatures prevail along shore, for more than two hundred miles (except in Todos Santos and San Quintin bays). The average surf temperatures at five "stations from the Mexican border northward to 1st and Coronado are always colder than the station at Scripps Pier. They are colder by average amounts of 1.2°, 0.5°, 0.9°, and 1.8° (C.) in the quarters beginning August, November, February and May respectively." (Dale F. Leipper, MS., based on 42, 33, 28 and 27 observations by Francis P. Shepard and associates.) Two distinct hydrographic areas are involved, with a rather sharp boundary approximately at the harbor entrance. Even slight changes might be expected to shift this boundary.

Under the conditions that existed from 1850 to 1870, when the temperatures seem to have been warmer, the cooler water to the southward probably extended to the harbor mouth and into the entrance at least as far as the point where the temperatures were taken. Changes since induced by man, particularly by the construction of a jetty, seem to confine the cooler waters so that they often do not penetrate the harbor entrance. In 1859, according to the chart of San Diego Bay, California made "From a Trigonometrical Survey under the direction of A. D. Bache Superintendent of the Survey of the Coast of the United States" (consulted in the Junipero Serra Museum), Zuñiga Shoal was tied to Point Loma by a bar having a depth of 21 to 22 feet at mean lower low water. A submerged spit extending from the shoal on the east side of Point Loma south of Ballast Point, in a direction somewhat east of south, and a slight eastward twist of the entrance channel inside the bar, both suggest that the flow in and out of the harbor involved the cooler water to the eastward and southward. If we make that assumption, and further assume that the temperature of that water was as much cooler than that at La Jolla (and elsewhere from off Point Loma northward) in 1854–60 as it was in 1946–47, when Shepard's observations were made, we may add the temperature differential determined by Leipper to the curve of the monthly means for 1854–60 (the heavy solid line in Fig. 6) to obtain a new curve that may well approximate the average along-shore temperatures from about 1850 to 1870 at points from Point Loma northward 50 miles or more. This curve, the dashed line in Fig. 6, may be regarded as a maximum estimate, and the true values may lie a little lower, but almost surely somewhat above the values that obtain at present from near Point Loma northward. Since various complications are involved, the preceding discussion should be regarded as leading merely to an approximate estimate.

By analogy and inference we may assume that the sea temperatures at Monterey were also high about 1850, or that the area of warm water extended close enough to Monterey to permit a spilling over of the warm-temperate fish fauna. About San Francisco, however, we must infer that no marked long-term changes in sea temperature have occurred since 1850.

The gradual upward trend of the air temperatures since about 1910, for all times of the year and for both San Diego and San Francisco (Figs. 1–3, 5), may well provide a false indication of warming sea temperatures. The trend might be taken as being in line with the "arctic amelioration of climate," but Dr. H. W. Ahlmann, the noted proponent of that concept, has indicated (personal communication) that the temperature changes may not have occurred (or may even have been reversed) in the approximate latitude of southern California. There is a small amount only of faunal evidence to favor a warming sea in this region over the last few decades. On analysis (suggested by Dale F. Leipper), it is found that the rise in mean air temperature at San Diego has been due almost entirely to a sharp rise in the daily maxima, which may be explained as due to the rapid growth of the city: taller buildings cut off the cooling sea breezes and more paving and masonry may cause a greater absorption of solar heat (Fig. 5). On the same basis we may perhaps account for the upward trend of the mean monthly and mean annual temperatures recorded for San Francisco (Figs. 2–3). When graphed by Leipper, the curves for the average daily maximum temperatures and for the population of San Diego correspond very closely in trend. The elevating effect of cities on temperature is a phenomenon recognized by meteorologists. For example, Hann (1903: 29–32; 1908: 38–39) stated that temperatures are usually increased 0.5 to 1.0° C. by cities. The sharp increase beginning with 1940 in the maximum daily temperatures at San Diego (beyond the dashed vertical line in Fig. 5) is definitely attributable to a change in the location of the weather station (p. 475). The trend of the surface sea temperature at La

Jolla has remained nearly level since the inception of the records in 1916, except for a midperiod rise due largely to the exceptionally warm years of 1926 and 1931, each of which was accompanied by a northward displacement of the coastwise fish fauna.

DISCUSSION AND SUMMARY

The general conclusion reached in this paper, that temperature plays a dominant role in the distribution of marine organisms, is in agreement with views generally held, for example by Ekman (1935) in his great monograph on the distribution of marine animals. The idea has been further stressed by a number of recent workers, including Stephenson (1941: 7) and Hutchins (1947). Organisms behave like thermometers in registering ocean temperatures. By reason of their stability they may often be of greater value in indicating average temperatures or in indicating peculiar seasonal distributions in temperature. Plants and animals are particularly valuable as indicators of local reversals in temperature changes with latitude.

The temporal fluctuations in temperature that seem to have been so frequent throughout the Quaternary have resulted in many faunal changes. Within historic time there have been some long-term as well as many single-season changes in temperature, with accompanying changes in the distribution of marine organisms. Such changes are stressed in the present paper, primarily for the fish faunas of the Pacific Coast of the United States. The 1850 and 1860 decades appear to have been in a prolonged warm period. The fish fauna at San Diego seems to have been much more nearly tropical than it now is, and the Monterey fauna seems to have comprised many elements now characteristic of southern California. The physical evidence for warmer waters is largely based on air temperature records. The strength of the faunal data involved is enhanced, when we note the fidelity with which the present fish faunas register the latitudinal variations in temperature. The northward occurrence of southern types in occasional years of high surface temperatures is particularly striking.

In correlating the changes in fish fauna with changes in the ocean temperatures I have given joint consideration to the biological and the physical data, in the belief that each may support the other. Consistency in the explanation of the natural phenomena of the sea has been my aim.

ACKNOWLEDGMENTS

For invaluable assistance in the interpretation and gathering of physical data I owe thanks to the master oceanographer, Harald U.

Sverdrup; to other colleagues at Scripps Institution of Oceanography, particularly Dale F. Leipper, George F. McEwen, Francis P. Shepard and Donald B. Sayner; to Dean Blake of the United States Weather Bureau, and to Margaret Calderwood of the Fish and Wildlife Service. Biological data and suggestions have come from E. Yale Dawson, now of the Allan Hancock Foundation of the University of Southern California; from my colleagues Martin W. Johnson and John D. Isaacs; from W. I. Follett, of the California Academy of Sciences; from George Y. Harry, Jr., and Donald L. McKernan, of the Oregon Fish Commission, and from John L. Hart, of the Pacific Biological Station, Fisheries Research Board of Canada. Historical aid has come from John Davidson, director of the Junipero Serra Museum of San Diego. Heber W. Blair has traced all the graphs for the figures. My wife, Laura C. Hubbs, has assisted in many ways.

LITERATURE CITED

Barnhart, P. S., and C. L. Hubbs
1946. Pontinus vaughani, a new scorpaenid fish from Baja California. Bull. Scripps Instn. Oceanogr., *5:* 371–390, fig. 1.
Carpenter, F. A.
1913. The climate and weather of San Diego California. San Diego Chamber of Commerce. xii + 118 pp., 28 figs.
Clemens, W. A., and G. V. Wilby
1946. Fishes of the Pacific Coast of Canada. Bull. Fish. Res. Bd. Canada, 68: 1–368, figs. 1–253.
Ekman, Sven
1935. Tiergeographie des Meeres. Akademische Verlagsgesellschaft M. B. H., Leipzig. xii + 542 pp., 244 figs.
Girard, Charles
1858. Fishes. In: General report upon the zoology of the several Pacific railroad routes. Explorations and surveys for a railroad route from the Mississippi River to the Pacific Ocean, *10* (4): 1–400, 21 pls.
Hann, Julius von
1903. Handbook of climatology. Part I. General climatology. Translated by Robert de Courcy Ward. The Macmillan Co., New York. xiv + 437 pp., 22 figs.
1908. Handbuch der Klimatologie. 3rd ed. I. Band. Allgemeine Klimalehre. J. Engelhorn, Stuttgart. xiv + 394 pp., 22 figs.
Hart, J. L.
In press. Increased abundance of an unusual British Columbia fish, the California pompano. Canad. Field Nat.
Hertlein, L. G., and U. S. Grant, IV
1944. The geology and paleontology of the marine Pliocene of San Diego, California. Part 1, Geology. Mem. San Diego Soc. nat. Hist., *2:* 1–72, figs. 1–6, 1 map, pls. 1–18.

HUBBS, C. L.

1946. Wandering of pink salmon and other salmonid fishes into southern California. Calif. Fish Game, *32:* 81–86.

HUBBS, C. L., AND ELIZABETH M. KAMPA

1946. The early stages (egg, prolarva and juvenile) and the classification of the California flyingfish. Copeia, 1946: 188–218, figs. 1–4, pl. 1.

HUBBS, C. L., AND R. R. MILLER

1948. The zoological evidence: Correlation between fish distribution and hydrographic history in the desert basins of western United States. In: The Great Basin, with emphasis on Glacial and Postglacial times. Bull. Univ. Utah, (Biol.) *10* (7): 17–165, figs. 10–29, map 1.

HUBBS, C. L., AND L. P. SCHULTZ

1929. The northward occurrence of southern forms along the Pacific Coast in 1926. Calif. Fish Game, *15:* 234–240, figs. 80–81.

HUTCHINS, L. W.

1947. The bases for temperature zonation in geographical distribution. Ecol. Monogr., *17:* 325–335, figs. 1–8.

JORDAN, D. S., AND C. H. GILBERT

1881a. List of the fishes of the Pacific Coast of the United States, with a table showing the distribution of the species. Proc. U. S. nat. Mus. (1880), *3:* 452–458.

1881b. Notes on the fishes of the Pacific Coast of the United States. Proc. U. S. nat. Mus., *4:* 29–70.

JORDAN, D. S., AND E. A. MCGREGOR

1925. Family Salmonidae (pp. 122–146, pls. V [pt.]–VIII). In: Record of fishes obtained by David Starr Jordan in Japan, 1922, by Jordan and Hubbs. Mem. Carneg. Mus., *10* (2): 93–346.

JORDAN, D. S., AND MASAMITSU OSHIMA

1919. Salmo formosanus, a new trout from the mountain streams of Formosa. Proc. Acad. nat. Sci. Philad. (1919), *71:* 122–124, 1 fig.

MCEWEN, G. F.

1916. Summary and interpretation of the hydrographical observations made by the Scripps Institution for Biological Research of the University of California 1908 to 1915. Univ. Calif. Publ. Zool., *15:* 255–356, figs. A-C, pls. 1–38.

SCHENCK, H. G., AND A. MYRA KEEN

1937. An index-method for comparing molluscan faunules. Proc. Amer. phil. Soc., *77:* 161–182, figs. 1–4.

SCHMITT, W. L.

1921. The marine decapod Crustacea of California. Univ. Calif. Publ. Zool., *23:* 1–470, figs. 1–165, pls. 1–50.

SMITH, J. P.

1919. Climatic relations of the Tertiary and Quaternary faunas of the California region. Proc. Calif. Acad. Sci., (4) *9:* 123–173, pl. 9.

STEINBECK, JOHN, AND E. F. RICKETTS

1941. Sea of Cortez. A leisurely journal of travel and research. The Viking Press, New York. x + 598 pp., 40 pls., 2 maps.

STEPHENSON, T. A.
1941. A summary account of the ecological survey of the South African coast carried out from the University of Cape Town during the years 1931–40. Privately printed by author: 1–14, figs. 1–2.

STEPHENSON, T. A., AND ANNE STEPHENSON
1948. Report of work done in North America in 1947–1948. Privately printed by authors: 1–11.

WALFORD, L. A.
1931. Northward occurrence of southern fish off San Pedro in 1931. Calif. Fish Game, *17:* 401–405, figs. 107–113.

PHYSICAL CHARACTERISTICS OF THE INSHORE ENVIRONMENT[1]

By

R. H. FLEMING
U. S. Navy Hydrographic Office
Washington, D. C.

INTRODUCTION

Oceanography, especially in the United States, has advanced very rapidly during the last two decades. This is true not only of the amount of data collected but also in the development of theories, the improvement of instrumentation, and in the training of investigators. At the same time there has been a tremendous increase of interest in the possible applications of oceanography to other fields of science and to various economic and military problems. Particularly in the phases of economic and military interest, attention is now focused upon the nearshore environment, a realm that most oceanographers have neglected in the past. In all sciences there is a strong inclination to go and graze in the far fields that look so green. Oceanographers, no less than others, have followed this lure and whenever opportunity has offered have turned their ships toward the deep sea. The relative simplicity of the great depths may be as responsible as their romantic appeal, because in many ways the problems of the inshore, shallow waters are far more complex than those of the open ocean.

Because of the increasing interest in the coastal regions, it is the purpose of this paper to discuss some of the characteristics of this environment and to indicate the means that must be employed in its investigation. The terms *inshore* and *nearshore* will be used synonymously without strict limitations of either depth or distance from shore. The terms do imply, however, that the regions under discussion are in some way or ways influenced by proximity to land and to shallow depths. By *physical characteristics*, all the physical, chemical, geological, and meteorological phenomena in or acting upon the water that influence the organisms living in the water and upon the sea bottom will be implied. As the enumeration of these phenomena alone would more than outrun the length of this discussion, only the broadest terms will be employed and the reader must supply the details.

---

[1] Presented at a symposium on "Ecology of Inshore Waters," September 10, 1948, sponsored jointly by The American Society of Limnology and Oceanography with the Ecological Society of America.

The inshore environment is, of course, unique in that it represents the region where the atmosphere, the ocean, and the underlying sea bottom mutually intersect. Because of the rise and fall of the sea surface caused by tide and wave action, there is truly a marginal zone that is by turn marine and terrestrial. The coastal realm therefore brings together the students of the land sciences and those of the sea. In certain regions, where there are almost completely landlocked bodies of water or where rivers enter the sea, the science of oceanography merges imperceptibly into that of limnology.

The inshore environment is complex, indeed, but the study of the physical characteristics and of the plants and animals found therein has been exceedingly fruitful. From the economic point of view, it is the region where most of the world's great fisheries are conducted. It is where organisms are not only plentiful but are most varied. It is the environment where most of the sedimentary rocks have been laid down. In order to understand the geologic history of the earth, it is essential that far more complete data be compiled upon the interrelations between the organisms, particularly potential fossil-forming types, and the physical characteristics of their environment.

It is obvious that the physical characteristics are the keys to both biological and geological problems. These factors can be used to characterize and compare the different environments, so that similarities and differences can be recognized and through the key organisms the terms that must be employed to describe the features of the oceans that existed in past eras.

THE PHYSICAL CHARACTERISTICS

It is not the purpose of this discussion to enumerate all the individual characteristics that make up the physical environment. However, a few of the more significant ones must be mentioned. First of all, because we are dealing with a marine environment, it is implied that the water contains dissolved salts in the relative proportions that characterize sea water. In regions of great dilution this relationship will obviously not hold in a strict sense, and there will be significant differences depending upon the origin of the diluting water. In this connection it should be pointed out that under certain circumstances runoff can be beneficial in that it will carry to the sea nutrient materials required by plant growth. At the other extreme, industrial wastes and sewage can be extremely harmful. The fresh water introduced by rivers and runoff from the land produces a variety of effects characteristic of the nearshore regions. Besides diluting the water, in the sense of concentration of dissolved materials, it carries with it suspended material of both organic and inorganic origin. This not only tends to reduce the transparency of the water but is, of course, the origin of much of the sediment accumulating on the bottom or being transported across the shelf to deeper water. The fact that the fresh water is lighter than the sea water tends to the formation of a dilute surface layer that overlies the more saline water. This may produce a variety of conditions. During spring in midlatitudes, the surface layer may be much colder than the underlying water. During the summer, the turbid surface layer may become very warm with an almost discontinuous decrease in temperature at the lower boundary of the dilute layer.

Motion of the water is generally more violent in the coastal regions than in the open sea. Currents that are so important in the transportation of organisms and their food are generally more rapid in the shallow water and in the proximity of land boundaries. Four types of water movement must be evaluated in describing an environment: permanent currents, wind-driven currents, tidal currents, and those currents associated with surface waves and river outflow. The rise and fall of the tide and the violent action of the surf on the shore are dominating features in the shape and character of the beach and shallow water zone. The transport of water, either along the beach or away from shore, caused by the onshore movement maintained by the waves, is also significant. All the various chemical constituents, particularly those affected by biological processes, such as the nutrient substances (nitrogen, phosphorus, and silicon), dissolved oxygen, and the carbon dioxide-calcium carbonate equilibrium, are also valuable parameters.

CHARACTERISTICS OF THE INSHORE ENVIRONMENT

Some of the physical characteristics have just been outlined. If it is possible to talk about "inshore" or "nearshore" environments as opposed to offshore or oceanic environments, there must be certain features that set them apart. The most obvious of these is the much greater range and variability of conditions near shore. First let us consider the regional variability. The highest temperatures found in the sea exist in land-locked waters in arid regions in low latitudes. In the Red Sea temperatures exceeding 40° C are common during the summer months, whereas temperatures exceeding 30° C are rare in the open ocean. Temperatures in tide pools and shallow water during the daylight hours rise rapidly and at night may fall considerably below those in adjacent deeper water. The opposite case of low temperatures does not apply so well, because when the temperature falls below $-2°$, ice is formed and sets the lower limit for temperature of the water. However, it is obvious that sea ice in shallow water will have more far-reaching effects upon the biological environment than sea-ice in deep water.

Salinity also shows a far greater range in coastal regions than in the open sea. To refer to the Red Sea once more, it is found that there salinities often exceed $40^{\circ}/_{oo}$, whereas in the open ocean they rarely are found to be higher than $37^{\circ}/_{oo}$. In land-locked areas and tide pools, of course, evaporation sometimes proceeds to the point where the salts are precipitated. The effects of dilution have already been mentioned, but it should be pointed out that in the open sea the salinities rarely drop below $33^{\circ}/_{oo}$.

The character of the bottom is also more variable, ranging all the way from solid rock to rapidly accumulating fine sediments off the mouths of such rivers as the Mississippi. The combined effects of waves and currents upon the organisms of the nearshore regions are certainly more violent than in the deep sea.

It is frequently stressed that the marine environment is extremely stable. From what I have just mentioned, it is obvious that inshore environments are in direct contrast by being extremely varied. Shallow coastal regions are also subject to a much greater range of conditions than the open sea. The rise and fall of the tide has little effect upon the deep sea environment, but it is certainly one of the important features in shallow water. This is true not only because of its effect upon the intertidal zone but also because of the tidal currents that sweep back and forth across the shelf. To thrive in the nearshore zone, organisms must be able to withstand a variable environment. They must be adapted to meet fluctuating temperatures, wide ranges in salinity, turbid water, shifting bottom sediments, and violent motions set up by waves and swell. At first glance, this picture looks rather dark for the organism, but it raises some very interesting questions. If the inshore environment is so variable, why does it appear to be so productive? Is it the proximity to land or is it the shoaler depths that make this great production possible? Answers to these questions and many others could undoubtedly be found in studies of carefully selected areas. How do areas with extensive shelves compare with coastal regions where there is deep water near shore? How does the productivity of an offshore area overlying an isolated bank compare with that of a shallow area near land? The recent discovery of the phantom bottom or deep scattering layer tends to indicate that in deep water the organisms may be relatively dilute but dispersed through a deep water column. The fact that shallow regions are essentially two-dimensional rather than three-dimensional may be very significant in their bioeconomy.

In order to characterize nearshore environments it will be necessary to set up certain questions that must be answered for each region. In many cases these will involve a comparison with the oceanic conditions in the adjacent areas. The boundaries will not be clear-cut but it will certainly reveal that the inshore environment, wherever it may be, is more variable than the nearby oceanic regions.

COMPARISON OF INSHORE ENVIRONMENTS

Since the physical environment is a complex of many variable factors, the only way in which different areas can be compared is by describing them according to some standard procedure. The list of characteristics should be held to a minimum and as our experience develops they should be arranged in some type of generic system that will permit classification. It appears that four general features might serve as the basis for such a classification of nearshore environments. These are:

a. The degree of isolation.
b. The type of coastline.
c. The latitude.
d. The relationship to adjacent land masses.

The degree of isolation is a qualitative expression indicating the extent to which the area under consideration is cut off from the adjacent sea areas. It is important because it reflects the amount by which conditions will differ from those in the open ocean. This can be illustrated by a single example. An area such as a fjord with a narrow and shallow entrance will be a very different environment than that in the adjacent open water.

The type of coast is believed to be a useful parameter because the topography of the sea floor near the coast is in general very similar to that of the land near the coast. This is the equivalent of saying that the location of the water line is irrelevant and that if sea level were lowered we would find essentially the same terrain extending outward to the edge of the shelf. The number of different types of coastline is a subject of debate, but one author recognizes eleven major types ranging from steep mountainous areas to river deltas and coral reefs. The reason for emphasizing the type of coastline is that it makes it possible to estimate many of the physical features of the regions from land maps and charts. These will be not only the general subsurface terrain but also the type of bottom, the amount of local runoff, the probable amount of suspended material and its character, exposure to storms, and the types of current.

Latitude is obviously a useful index to a number of features. Not only is it an index to the seasonal range and value of temperatures, but also it fixes the duration of seasons, intensity of illumination, and length of day.

Proximity to adjacent land masses is another qualitative indicator. Whether we are concerned with an area adjacent to a continent, a lone island, or part of an archipelago is obviously important. We must also know whether the area is on the east or west side of a continent because only then will we be able to predict the general type of circulation, the importance of upwelling, the type of tides, and certain other features.

Any system of classification is not of particular value without detailed information on conditions in the individual areas. Because extremely valuable conclusions could be drawn from detailed comparisons of different areas, it is believed that a program for systematic compilation of such studies should be undertaken. To be most useful they should consist of three general parts:

1. General average conditions.
2. Historical compilation.
3. Factors controlling the environment.

The first section should present as briefly as possible by means of graphs, charts, and tables the average "climatic" picture that would serve as the basis for comparison with similar analyses for other areas. All studies should be based upon the same outline and should employ standardized means of presentation so that comparisons may be made rapidly and easily. The historical compilation should contain all available types of records pertinent to the area for which they have been accumulated over a period of years. In most locations these data will be scanty and will be limited to such items as sea level data from tidal records, surface temperatures, surface salinities, river discharge values, and such records as may be available from meteorological offices. This information would be most useful in establishing empirical correlations between catch records, success of year classes, abnormal biological conditions, and certain physical features. The third section should be a discussion of the area, pointing out the factors that seem most important in controlling the local environment and indicating possible means by which variable factors might be predicted in advance.

In the foregoing no mention has been made of the biological information that might be included in such systematic studies. It is clear that compilations of the type just outlined would be of the greatest help to biologists and fisheries experts and would provide the type of information that is urgently needed by those concerned with paleoecology. There is another use, however, that may not be quite so obvious. The physical environment, in very large part, determines the feasibility of fishing, the types of equipment to be employed, and the size of vessel necessary to operate under a given set of conditions. From the systematic analyses just outlined, it is easy to see how such information could be used to great advantage in planning fishery development in hitherto unexploited regions and in improving techniques in regions where fisheries have developed by traditional trial and error methods.

SUMMARY AND CONCLUSIONS

As the inshore environment is always so variable, it is important that methods of investigation be devised that make field investigations as rapid and simple as possible. Because of the importance of comparative studies it is likewise important that methods of study and reporting be standardized. This requires careful planning and coordination. By classification and systematized descriptions of inshore environments information would be available that would serve not only the biologist, paleontologist, geologist, and fisheries investigator but would be of value in fishery development and the expansion into new types of fisheries and into new areas.

DIATOM POPULATIONS ASSOCIATED WITH EDDIES OFF SOUTHERN CALIFORNIA IN 1941[1]

By

MARSTON C. SARGENT AND THEODORE J. WALKER

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

During 1941 the E. W. Scripps made ten cruises off southern California, each covering the same area of about 80,000 square kilometers. The cruise plan consisted of a grid of approximately 40 stations. At each station a catch was made with the Allen five-liter closing bottle every ten meters from the surface to 60 meters. The catch was filtered through a net of No. 25 silk bolting cloth, and the diatoms so collected were later identified, counted and recorded as the number of diatoms per liter. The laboratory work was done by the late W. E. Allen, who had no opportunity to prepare more than a general report, which is now being edited for publication. For the purpose of this paper the number of diatoms per liter is taken as representing the number per square centimeter from the level at which the catch was made to the level of the next catch ten meters below. The sum of the seven catches at a station represents, according to this convention, the diatom population per square centimeter to a depth of 70 meters. Collections were not made at greater depths. However, the progressive reduction of the population down to 60 meters suggests that the diatoms below this depth would not change the totals sufficiently to affect the validity of the arguments presented here.

At each station, observations were made of temperature, salinity, and oxygen to a depth of 500 meters. Dynamic height anomalies (Sverdrup, 1947) are shown on all charts of this report as an indication of the character of the surface water and its horizontal movements. The general dynamic features of these waters have been discussed by Sverdrup and Fleming (1941), who showed that in the spring and early summer two types of surface water (down to about 100 meters) are present in this area: (1) "new" water which has been recently drawn to the surface by upwelling and (2) "old" water which has been at the surface for a relatively long time. New water is characterized by low temperature, high salinity, and a shallow thermocline. At this time of year there is typically in the western part of the cruise area and beyond it a south-flowing current of old water. Along its left-hand boundary one or more cyclonic eddies of new water appear, and inshore from these there may be a current of old water flowing northwesterly. Sverdrup and Fleming discussed the evidence for the theory that the two oppositely-directed currents represent two sides of a single large eddy. The pattern of circulation outlined above becomes less evident as upwelling of new water diminishes later in the year.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 386.

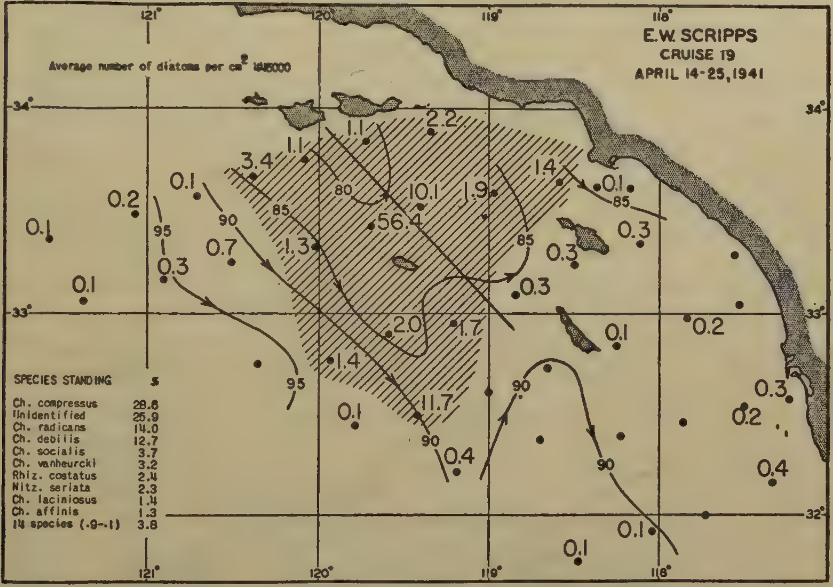
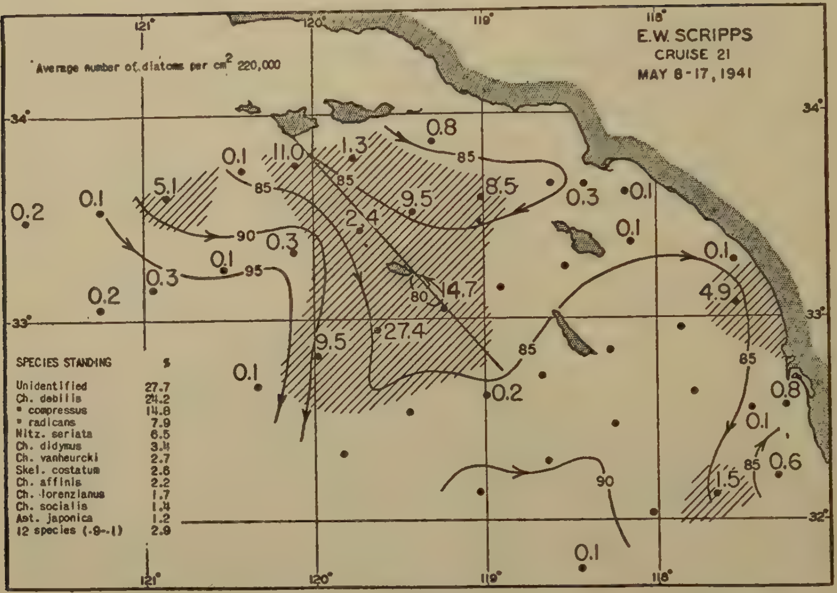

E.W. SCRIPPS
CRUISE 19
APRIL 14-25,1941
Average number of diatoms per cm² 146000
121°
120°
119°
118°
34°
33°
32°
SPECIES STANDING %
Ch. compressus 28.6
Unidentified 25.9
Ch. radicans 14.0
Ch. debilis 12.7
Ch. socialis 3.7
Ch. vanheurcki 3.2
Rhiz. costatus 2.4
Nitz. seriata 2.3
Ch. laciniosus 1.4
Ch. affinis 1.3
14 species (.9-.1) 3.8
2.2
1.1
1.1
3.4
14
0.1
0.1
0.2
0.1
0.1
80
85
10.1
1.9
85
56.4
95
90
0.7
1.3
0.3
85
0.3
0.3
0.3
2.0
1.7
0.1
0.2
1.4
95
90
0.3
0.2
0.1
11.7
90
0.4
90
0.4
0.1
0.1

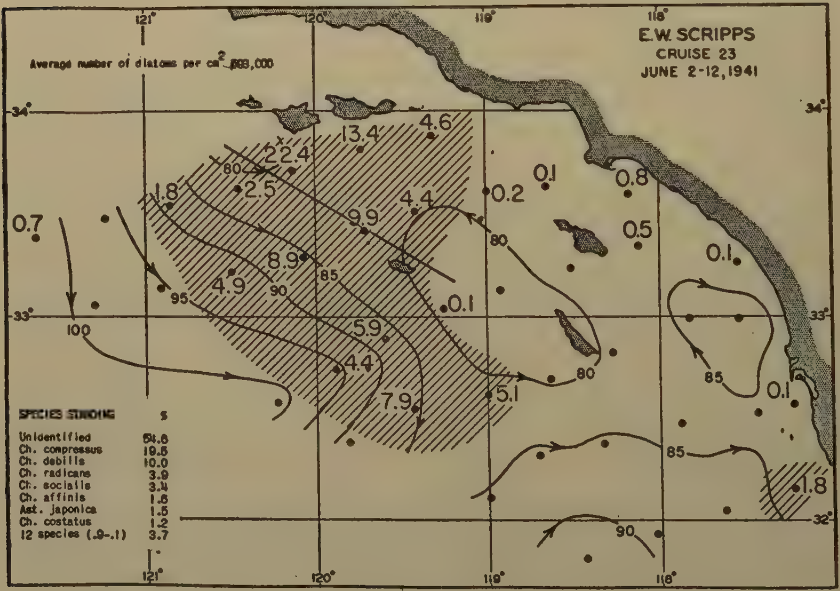
Figure 1. Distribution of diatoms and hydrography, Cruise 19. At each station (indicated by a dot), phytoplankton catches were made every ten meters from the surface to 60 meters. Figures at stations show percentage contribution of each station to cruise total of diatoms. Hatched area includes all stations contributing at least one per cent to cruise total. Numbered and arrowed contours show topography of sea surface in dynamic centimeters relative to 500 decibar surface. Arrows indicate inferred direction of flow of surface currents. Close spacing of contours indicates high current velocities (maximum velocities 8 miles per day).

Michael (1921) showed, for a line of stations extending 125 miles out from Point Fermin, a negative correlation between the population density of the phytoplankton and the temperature of the surrounding water, indicating that large populations are, on the average, found in


E.W. SCRIPPS
CRUISE 21
MAY 8-17, 1941
Average number of diatoms per cm² 220,000
121°
120°
119°
118°
34°
33°
32°
0.8
85
0.1
11.0
1.3
85
5.1
85
9.5
8.5
0.3
0.1
0.2
0.1
2.4
90
0.3
0.1
0.1
95
0.1
0.3
80
14.7
4.9
0.2
85
33°
9.5
27.4
85
0.1
0.2
0.8
0.1
85
0.6
1.5
90
0.1

SPECIES STANDING	%
Unidentified	27.7
Ch. debilis	24.2
" compressus	14.8
" radicans	7.9
Nitz. seriata	6.5
Ch. didymus	3.4
Ch. vanheurcki	2.7
Skel. costatum	2.6
Ch. affinis	2.2
Ch. lorenzianus	1.7
Ch. socialis	1.4
Ast. japonica	1.2
12 species (.9-.1)	2.9

Figure 2. Distribution of diatoms, Cruise 21. See Fig. 1 for description.


E.W. SCRIPPS
CRUISE 23
JUNE 2-12, 1941
Average number of diatoms per cm² 688,000
121°
120°
119°
118°
34°
33°
32°
4.6
13.4
22.4
80
2.5
1.8
0.1
0.2
0.8
4.4
0.7
9.9
0.5
80
8.9
85
4.9
90
0.1
95
0.1
100
5.9
4.4
80
85
7.9
5.1
0.1
85
1.8
90

SPECIES STANDING	%
Unidentified	54.6
Ch. compressus	19.5
Ch. debilis	10.0
Ch. radicans	3.9
Ch. socialis	3.4
Ch. affinis	1.6
Ast. japonica	1.5
Ch. costatus	1.2
12 species (.9-.1)	3.7

Figure 3. Distribution of diatoms, Cruise 23. See Fig. 1 for description.

recently upwelled water. Sverdrup and Allen (1939) found that large diatom populations in southern California waters were limited to new water, whereas water which had spent a long time on the surface was always characterized by a small diatom population.

OBSERVATIONS

For four cruises in 1941 Figs. 1 to 4 show the percentage contribution of each station to the cruise total of all diatoms. It is clear that the distribution of the diatom population is orderly and that the spacing of the stations is close enough to indicate the population at all points. In each figure, the area embracing all stations contributing at least 1% to the cruise total has been hatched. For convenience, the outline of this area will be called the one-per-cent contour.[2]

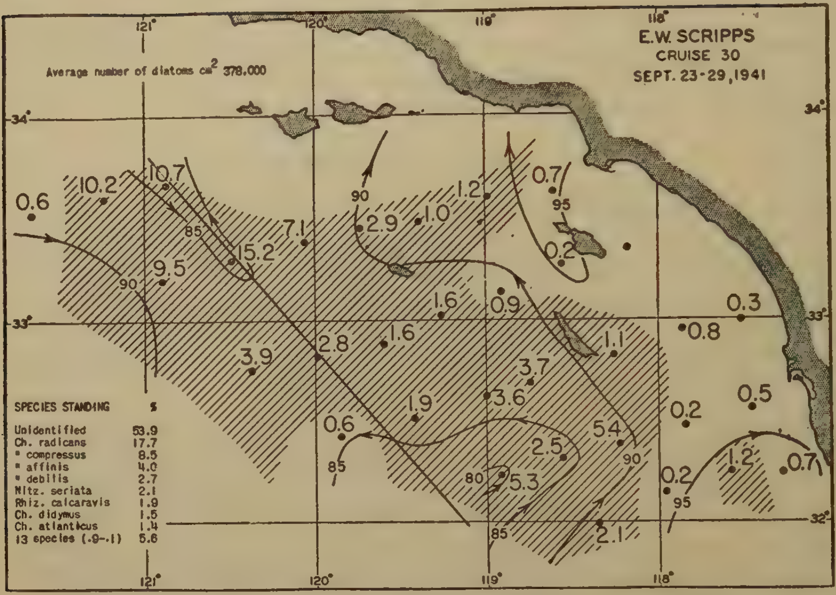
TABLE I. Distribution of Diatoms off Southern California in 1941

Cruise	*Month*	*Av. No. of diatoms per station (thousands per cm²)*	*No. of stations occupied*	*No. of stations with 1% or more of total population*	*% of total population at these stations*
17	March	1020	41	21	95.9
19	April	445	41	13	95.7
21	May	220	42	11	95.8
23	June	593	40	13	97.9
25	June	493	40	10	96.4
27	July	51	41	21	96.2
28	August	500	38	15	94.0
30	September	378	33	21	96.0
32	October	131	38	23	97.9
33	November	8	33	15	94.3

As shown on Table I, this contour encloses more than 90% of all the diatoms observed on each cruise. Fig. 5 shows that the use of either the half-per-cent or the two-per-cent contour would not greatly change the proportion of the population enclosed. For instance, the half-per-cent contour would enclose an additional 39 stations (four

---

[2] This outline can be defined precisely (on the assumption that the number of diatoms changes gradually from station to station) as the locus of all points such that if a station had been occupied at any one of them it would have contributed one per cent to the cruise total. The cruise total is the sum of the station populations. The station population has already been defined as the population to 70 meters per cm² of surface. The cruise total multiplied by the average area represented by one station (about 2,000 km²) equals the total population in the whole area covered by the cruise.


121°
120°
119°
118°
34°
33°
32°
Average number of diatoms cm^2 378,000
E.W. SCRIPPS
CRUISE 30
SEPT. 23-29, 1941
0.6
10.2
10.7
85
15.2
7.1
90
2.9
1.0
1.2
0.7
95
0.2
90
9.5
1.6
0.9
0.3
0.8
3.9
2.8
1.6
1.1
3.7
3.6
1.9
0.6
85
0.2
0.5
5.4
2.5
90
80
5.3
0.2
95
1.2
0.7
85
2.1

SPECIES STANDING	%
Unidentified	53.9
Ch. radicans	17.7
" compressus	8.5
" affinis	4.0
" debilis	2.7
Nitz. seriata	2.1
Rhiz. calcaravis	1.9
Ch. didymus	1.5
Ch. atlanticus	1.4
13 species (.9-.1)	5.6

Figure 4. Distribution of diatoms, Cruise 30. See Fig. 1 for description.

stations per cruise) and would increase the enclosed population by about 5% per cruise. On the other hand, use of the two-per-cent contour would exclude 49 stations (five per cruise) included by the one-per-cent contour and would diminish the enclosed population by about 6% per cruise. Neither alternate line would affect the conclusion that nearly all of the diatom population during these cruises was found regularly in a single sharply defined area of limited extent.

During most of the 1941 cruises, particularly in the spring and early summer months, a moderately intense cyclonic eddy was present in the northern part of the cruise area. In Figs. 1 to 4, lines have been drawn through the centers of such eddies, in a direction parallel to the flow of the offshore current, or if this could not be determined, in a southeasterly direction. In most cruises these lines divided the populations in such a way that the larger part was on the offshore (southwest side) of the line. Except during cruises 27, 32, and possibly 30, over half of all the diatoms collected occurred in these very restricted patches (Table IIA). In the exceptional cruises the eddies were close to the outermost line of stations, so that it is likely that the area of densest population was not sampled at all. Thus, it appears that the cyclonic eddies in the northern parts of the area characteristically contained large diatom populations which were especially dense in the

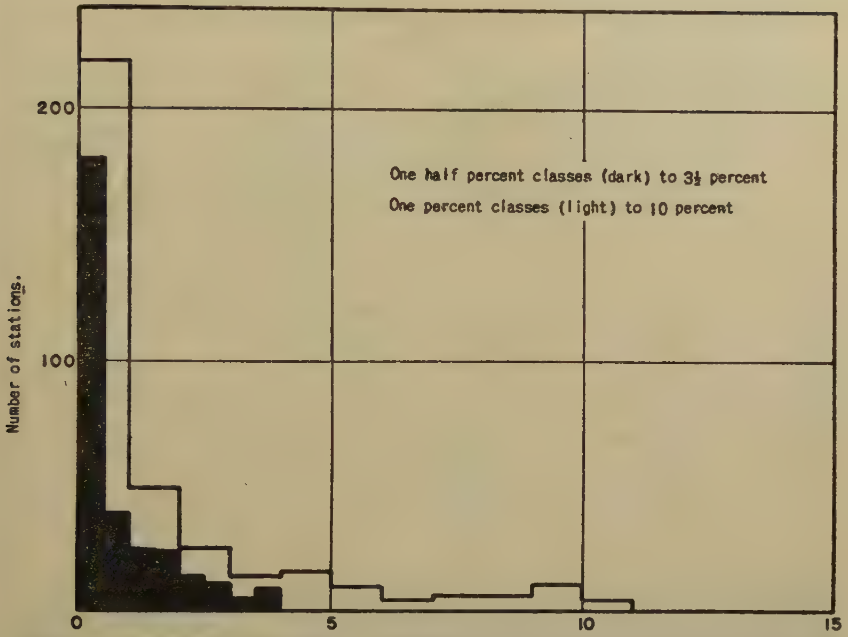


Figure 5. Frequency distribution of contributions of individual stations to cruise totals of all diatoms. The shaded histogram shows stations contributing less than 4%. The open histogram shows stations contributing less than 11%. Ten stations, each contributing more than 11%, are not included.

offshore half of the eddies (Table IIB). As this water flowed southeastward the populations developed and then were scattered or destroyed. Whether the greatest development was reached before or after the water entered the cruise area from the northwest is unknown, but the rapid decay and almost total disappearance of the population within the local waters is shown by the paucity of diatoms just outside of the eastern, southern and western boundary to the productive area.

This pattern of population densities remained approximately constant throughout the year, although the total population varied from cruise to cruise (Table I) and although the species of diatoms involved varied in relative abundance from season to season. During several cruises a single species accounted for more than a quarter of the entire population, and in every cruise a dozen species or less comprised from one-third to nine-tenths of the population. Seasonal succession and other changes in the composition of the list of leading species shown on the charts (Figs. 1 to 4) are matters to be discussed at length elsewhere.

TABLE II. DISTRIBUTION OF DIATOMS IN CYCLONIC EDDIES OFF SOUTHERN CALIFORNIA

	A		B	
Cruise	*No. of stations inside 1% contour and in southwestern half of cyclonic eddy*	*% of total population at these stations*	*Av. % of total population at each station inside 1% contour*	*Av. % of total population at each station inside 1% contour and in southwestern half of eddy*
17	10	64.1	4.6	6.4
19	8	79.0	7.4	9.9
21	4–5	50.3–65.0	8.7	12.5
23	9–10	51.3–73.7	7.5	5.7
25	7	80.5	9.6	11.5
27	8	16.4	4.6	2.1
28	7–8	51.3–60.8	6.3	7.3
30	3–6	23.6–52.3	4.6	7.9
32	8	37.3	4.2	4.7
33	11	81.1	6.3	7.4

The point to be made here is that, although the composition of the population changed, the pattern as a whole bore a remarkable constant relationship to the hydrography.

In the present report we will describe in detail only the observations made during the spring months when conditions were most clearcut and uniform from cruise to cruise. Therefore, discussion will be limited to those observations, and especially to those of cruise 19, which was representative of the spring cruises.

A number of the individual species observed during cruise 19 were distributed in patterns as clearly related to the hydrography as that of the diatom population as a whole. Figs. 6 to 11 show the distribution of six separate species during this cruise. As in previous charts, the hatching covers those stations contributing at least one per cent to the cruise total for the species. It is clear that *Nitzschia seriata* (Fig. 6) and *Chaetoceros radicans* (Fig. 7) had essentially the same distribution as the whole diatom population, as did *C. lorenzianus*, *N. longissima*, and several minor species. The boundaries of the area of high population were not determined, then, by a single preponderant species. As suggested by Michael, a whole group of species appears to be controlled in the same way by the same set of factors.

A second type of distribution is that of *C. socialis* (Figs. 10, 14). This species and *C. compressus* (Fig. 11), *C. debilis* (Fig. 15) and *C. didymus* were very similar in being restricted essentially to a few sta-

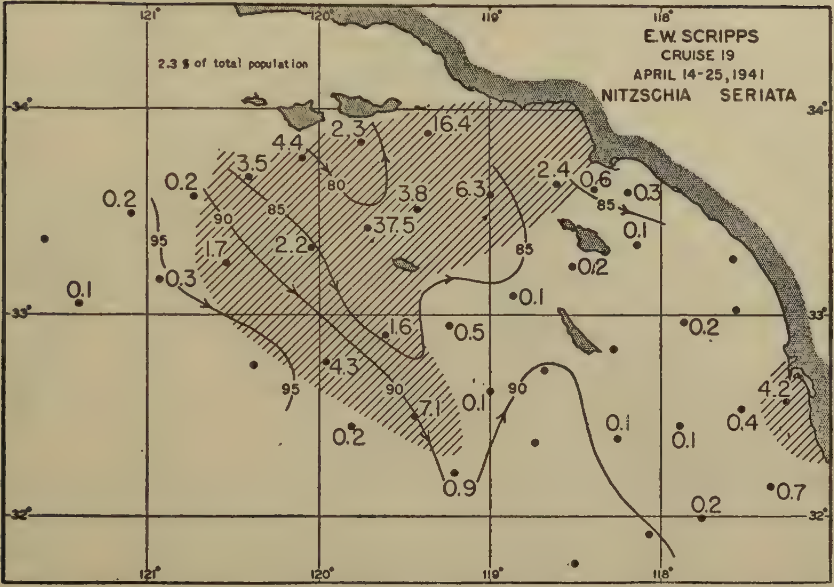

E.W. SCRIPPS
CRUISE 19
APRIL 14-25, 1941
NITZSCHIA SERIATA
2.3 % of total population
121°
120°
119°
118°
34°
33°
32°
0.2
0.1
0.2
0.3
95
90
85
80
3.5
4.4
2.3
16.4
1.7
2.2
37.5
3.8
6.3
2.4
0.6
0.3
0.1
0.2
0.1
1.6
0.5
4.3
7.1
0.1
0.2
0.9
0.1
0.1
0.4
4.2
0.2
0.7

Figure 6. Distribution of *Nitzschia seriata*, Cruise 19. At each station (indicated by a dot), phytoplankton catches were made every ten meters from the surface to 60 meters. Figures at stations show percentage contribution of each station to cruise total of the species. Hatched area includes all stations contributing at least one per cent to cruise total.

tions within the one-per-cent contour. Of these four species, all but *C. didymus* represented important fractions of the total diatom population, as shown in Fig. 1. At station 24 (Lat. 33° 25′ N, Long. 119° 43′ W) more than two-thirds of the cruise total for each of these species was collected, and they were also abundant at station 46 (Lat. 32° 29′ N, Long. 119° 26′ W). These species were common only in the neighborhood of these two stations, and in general they were not observed at all outside of the one-per-cent contour for all diatoms.

A radically different kind of distribution is that of *Rhizosolenia alata* (Fig. 8). The distribution of this species, as well as *Dactyliosolen antarcticus* (Fig. 9), *Rhizosolenia hebetata, Chaetoceros dadayi, C. atlanticus* and *C. messanensis*, is complementary to that of the main group. On the whole, diatoms of these six species were few or absent at populous stations but were relatively numerous at many stations outside the populous areas.

Thus, many of the species collected during cruise 19 can be divided into groups according to their distribution as follows:

1. Common, occurring at most stations in the eddy but rarely elsewhere. *Chaetoceros radicans*, etc.

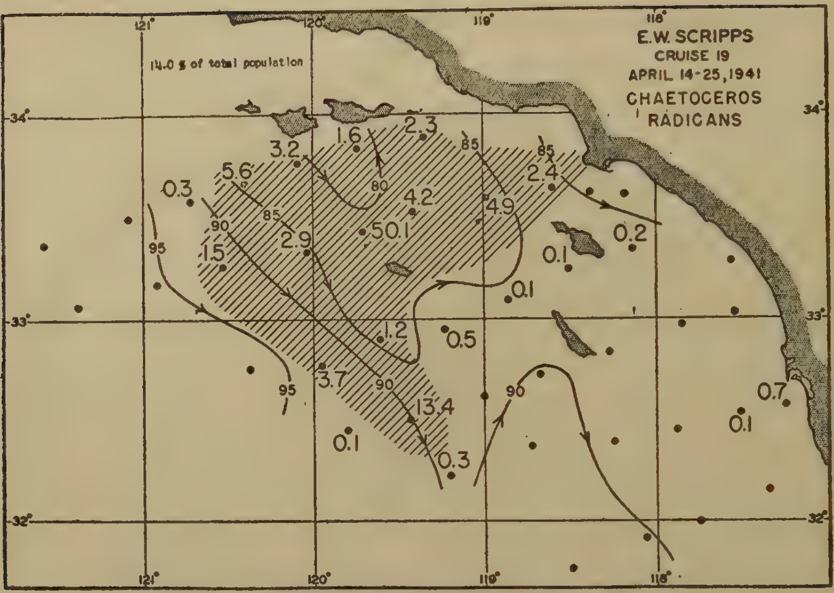

121°
120°
119°
118°
34°
33°
32°
E.W. SCRIPPS
CRUISE 19
APRIL 14-25,1941
CHAETOCEROS
RADICANS
14.0 % of total population
1.6
2.3
3.2
5.6
0.3
85
80
4.2
2.4
4.9
90
50.1
0.2
95
1.5
2.9
0.1
0.1
1.2
0.5
3.7
90
13.4
0.7
0.1
0.1
0.3

Figure 7. Distribution of *Chaetoceros radicans*, Cruise 19.

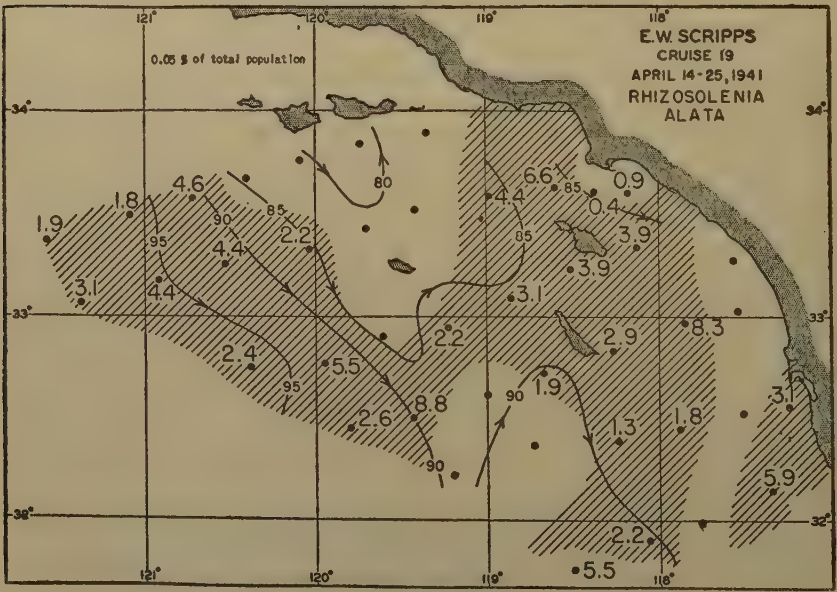

121°
120°
119°
118°
34°
33°
32°
E.W. SCRIPPS
CRUISE 19
APRIL 14-25,1941
RHIZOSOLENIA
ALATA
0.05 % of total population
80
4.6
1.8
1.9
85
90
95
6.6
4.4
0.9
0.4
2.2
4.4
3.9
3.9
3.1
4.4
3.1
2.2
8.3
2.9
2.4
5.5
1.9
95
2.6
8.8
1.3
1.8
3.1
90
5.9
2.2
5.5

Figure 8. Distribution of *Rhizosolenia alata*, Cruise 19.

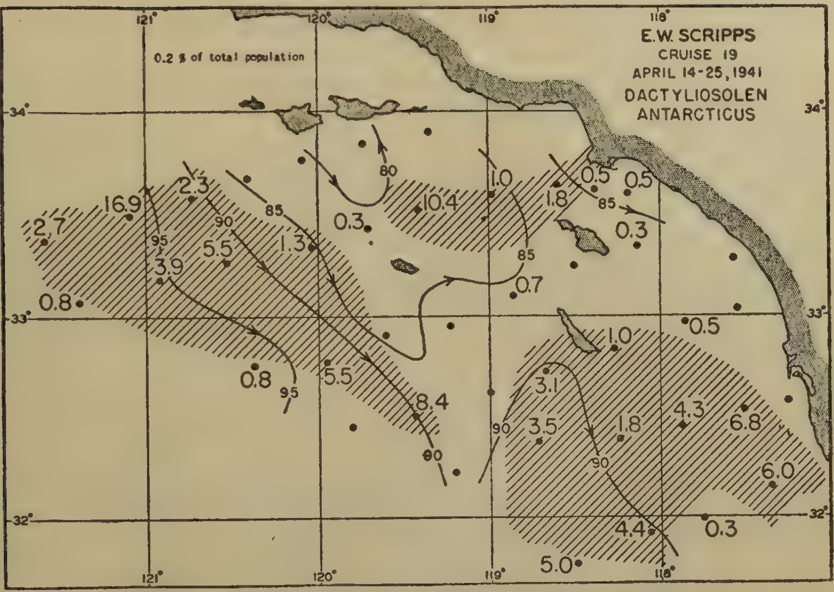

E.W. SCRIPPS
CRUISE 19
APRIL 14-25, 1941
DACTYLIOSOLEN
ANTARCTICUS
0.2 % of total population
121°
120°
119°
118°
34°
33°
32°
2.7
16.9
2.3
0.8
3.9
5.5
1.3
0.3
10.4
1.0
1.8
0.5
0.5
0.3
0.7
0.8
5.5
8.4
1.0
0.5
3.1
3.5
1.8
4.3
6.8
6.0
4.4
0.3
5.0
80
85
90
95

Figure 9. Distribution of *Dactyliosolen antarcticus*, Cruise 19.

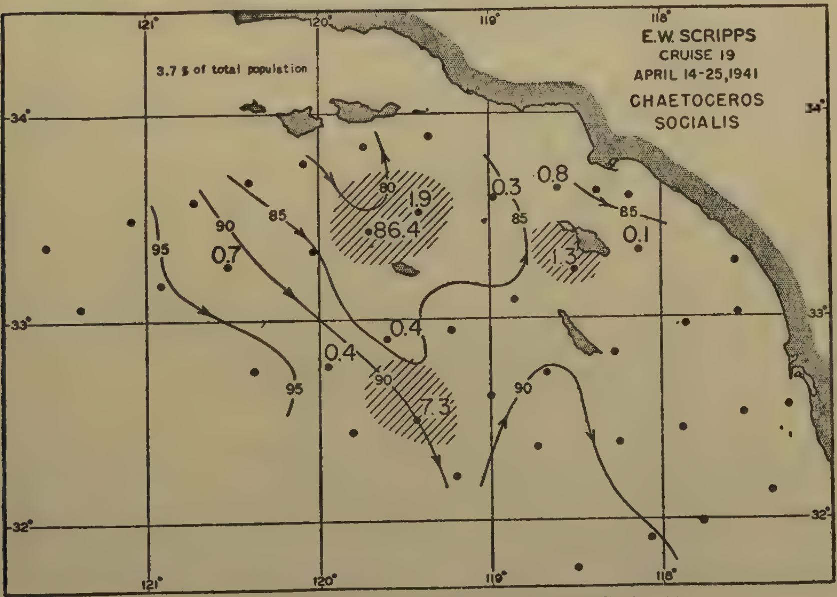

E.W. SCRIPPS
CRUISE 19
APRIL 14-25, 1941
CHAETOCEROS
SOCIALIS
3.7 % of total population
121°
120°
119°
118°
34°
33°
32°
1.9
86.4
0.3
0.8
0.1
1.3
0.7
0.4
0.4
7.3
80
85
90
95

Figure 10. Distribution of *Chaetoceros socialis*, Cruise 19.

2. Restricted, occurring at only a few stations in the cyclonic eddy. *Chaetoceros socialis*, etc.
3. Complementary, occurring seldom within the eddy, but generally elsewhere. *Rhizosolenia alata*, etc.

Species intermediate in character between those of groups 1 and 2 are to be expected, and it is not unlikely that an extensive study would show that the distribution of some species changes with, for instance, the seasons. However, all species listed above retained their characteristic distribution throughout the spring cruises of 1941. Many other species observed are not readily fitted into this scheme. In particular, some were so scarce that it is only possible to conclude that they did poorly over the entire cruise area.

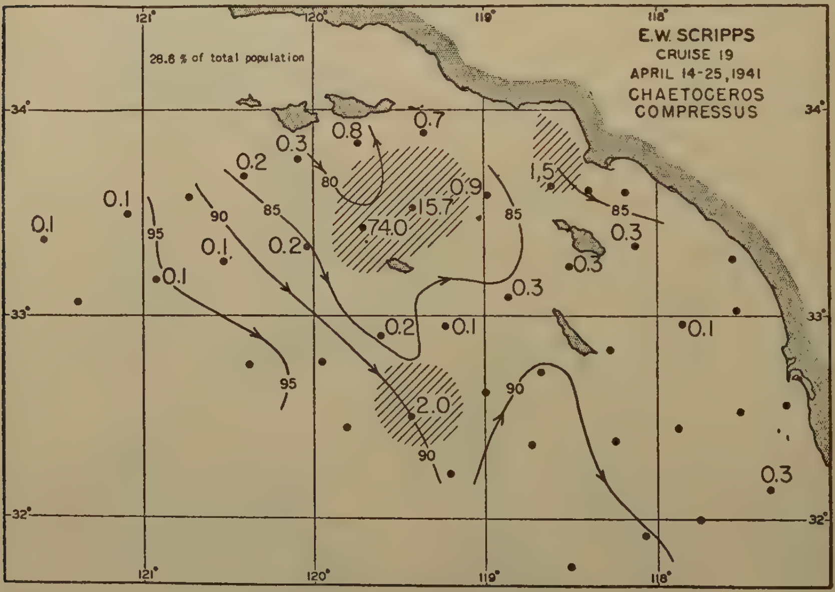

E.W. SCRIPPS
CRUISE 19
APRIL 14-25, 1941
CHAETOCEROS COMPRESSUS
28.6 % of total population
121°
120°
119°
118°
34°
33°
32°
74.0
15.7
2.0
1.5
0.9
0.8
0.7
0.3
0.2
0.1
80
85
90
95

Figure 11. Distribution of *Chaetoceros compressus*, Cruise 19.

If data for other years substantiate this division, each of the groups can be considered as belonging to a "plankton element" defined by Gran (1902: 74 ff) as an assemblage having similar ecological requirements. Since Gran's paper, many such ecological requirements have been discovered and elucidated, particularly by Harvey (1947, and papers referred to therein).

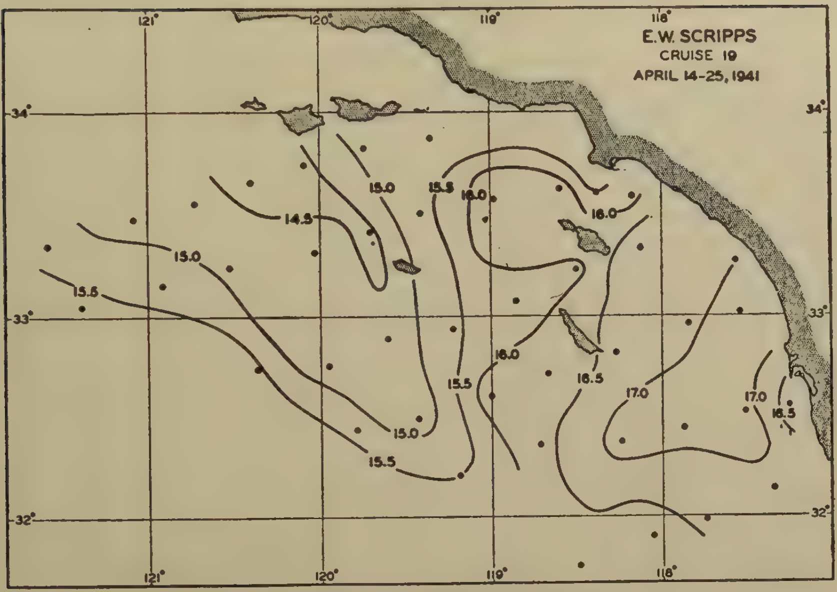

121°
120°
119°
118°
E.W. SCRIPPS
CRUISE 19
APRIL 14-25, 1941
34°
33°
32°
15.0
15.5
16.0
14.5
15.0
15.5
16.0
16.0
15.5
15.0
15.5
16.5
17.0
17.0
16.5
121°
120°
119°
118°

Figure 12. Surface temperatures, Cruise 19.

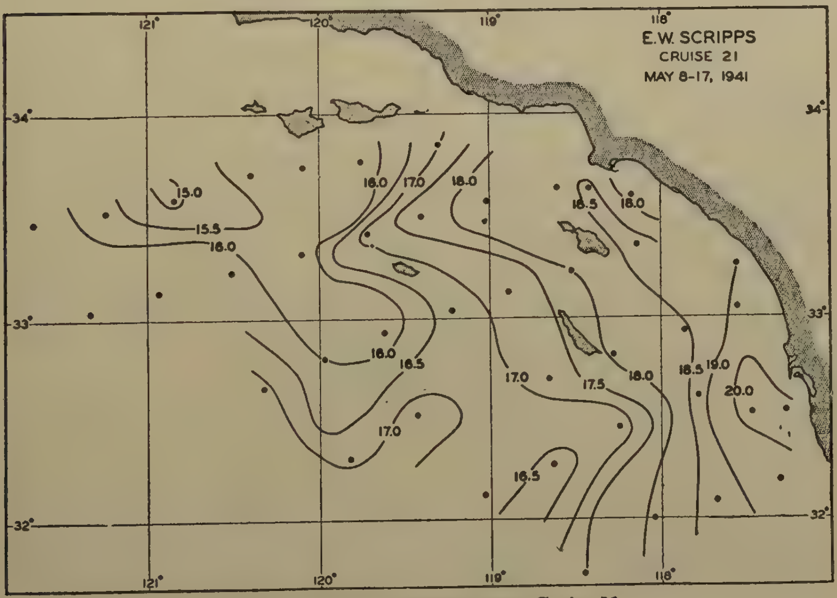

121°
120°
119°
118°
E.W. SCRIPPS
CRUISE 21
MAY 8-17, 1941
34°
33°
32°
15.0
15.5
16.0
16.0
17.0
18.0
16.5
18.0
16.0
16.5
17.0
17.0
17.5
18.0
18.5
19.0
20.0
16.5
121°
120°
119°
118°

Figure 13. Surface temperatures, Cruise 21.

INTERPRETATION

A close relationship existed between these groups of plankton and the circulation features of the water. It is inconceivable that the small differences in temperature observed, or possible small differences in average insolation, could account for the distribution of the diatoms. The depth of the mixed layer, which increased with distance from shore, as Sverdrup and Fleming observed, could not by itself account for the distribution. Rather, it is necessary to look for a complex of physical, chemical and biotic factors whose individual values and combined effect varied from day to day and from point to point. The over-all effect is that freshly upwelled water is more fertile than old water, as pointed out by Michael (1921) and by Moberg (1928). The details of the distribution of the diatoms during the spring of 1941 can be related to this phenomenon as follows.

Freshly upwelled water entered the area of observations from the northwest intermittently, and its concentration in some areas was higher than elsewhere as indicated by low temperatures in the centers of the cylonic eddies (Figs. 12, 13). In these areas, as at stations 24 and 46 in cruise 19, large populations occurred, with species in groups one and two predominating. From the fact that the very densely populated stations, which were also the only stations at which group-two diatoms were found in appreciable numbers, often occurred singly or in pairs, it can be inferred that the diameter of the areas of highest concentration of upwelled water were close to the station grid interval of 20–30 kilometers. Surrounding these areas was a larger one in which the concentration of new water diminished with increasing distance from the centers of population, as indicated by higher temperatures and lower salinities. This area supported a large population in which the species of group one were conspicuous, while diatoms of group two were rare or absent. The remainder of the area covered by the cruises was, on the whole, occupied by water which was of high temperature and low salinity and therefore had probably been on the surface for some time. The diatom population was small, but species of group three were found in their greatest abundance here.

The diminution in population density was probably controlled by different combinations of factors in different parts of the eddy system. Toward the east, that portion of the upwelled water which was carried around the eddy, except near its periphery, had little opportunity for mixing with old water, and the diminution of the population in the inshore half of the eddy must have resulted from processes occurring within the water body, such as exhaustion of nutrients, increase of grazing, etc., which produced combined effects as shown by Braarud

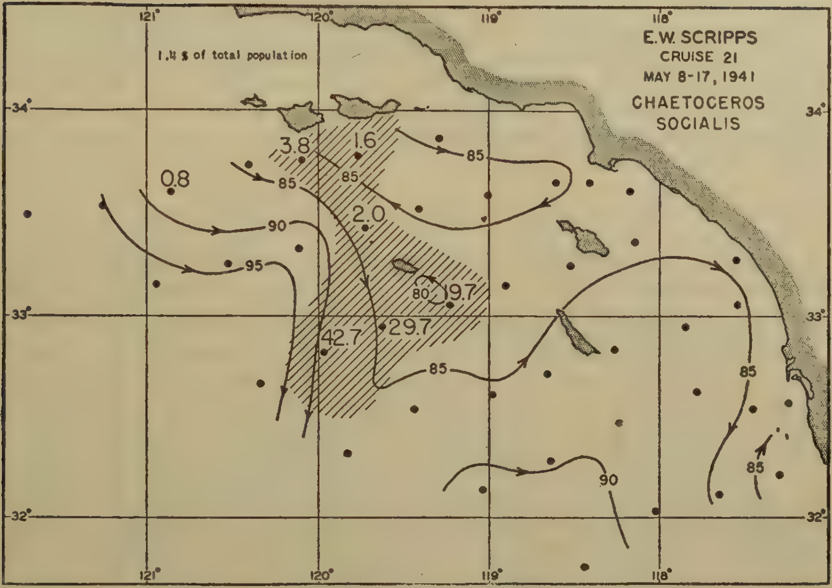

121°
120°
119°
118°
1.4 % of total population
E.W. SCRIPPS
CRUISE 21
MAY 8-17, 1941
CHAETOCEROS
SOCIALIS
34°
33°
32°
3.8
1.6
0.8
85
90
95
2.0
80
19.7
42.7
29.7
85
90

Figure 14. Distribution of *Chaetoceros socialis*, Cruise 21.

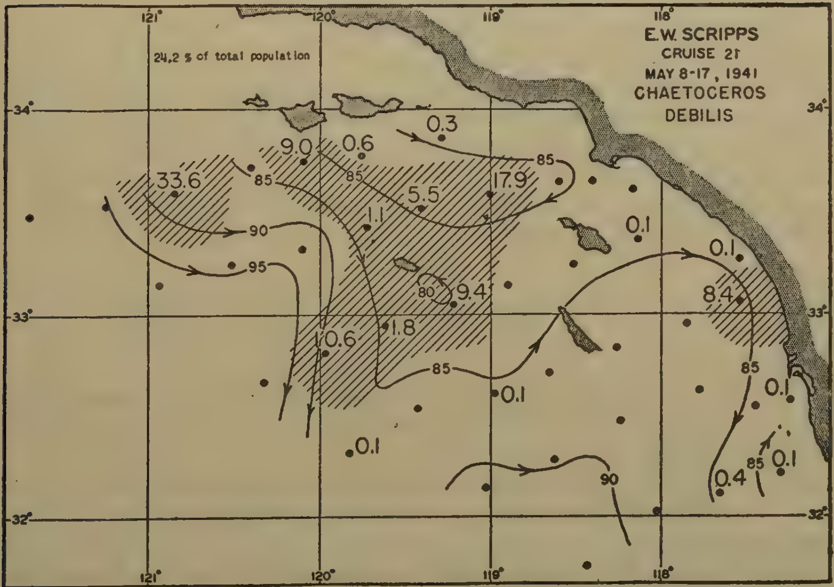

121°
120°
119°
118°
24.2 % of total population
E.W. SCRIPPS
CRUISE 21
MAY 8-17, 1941
CHAETOCEROS
DEBILIS
34°
33°
32°
0.3
9.0
0.6
33.6
85
17.9
5.5
1.1
90
95
0.1
0.1
8.4
80
9.4
0.6
1.8
0.1
0.1
0.1
0.4
0.1
0.1
90

Figure 15. Distribution of *Chaetoceros debilis*, Cruise 21.

(1935) and Fleming (1939). To the west, intense lateral mixing with offshore water which was both infertile and barren (Sverdrup and Fleming, 1941: 331) would reduce both the concentration of nutrients and the density of the population, and, in combination with the other processes, produce the steep population gradient on the offshore side of the eddies.

SUMMARY

In waters off southern California in 1941, the diatom population was concentrated at any one time in one or a few relatively small patches. These patches were closely related to cyclonic eddies of freshly upwelled water entering the area of observation from the north (particularly in the spring when upwelling was strongest). The population was most abundant in the offshore halves of the eddies where the upwelled water, judged by its low temperature and estimated path of travel, was most concentrated and had been present on the surface the shortest time. Populations in water which had apparently travelled some distance around the eddies were less abundant, presumably because of exhaustion of nutrients, grazing, and other biological processes. Abundance diminished more sharply to the west than to the east because, in addition, there was intense lateral mixing with barren infertile offshore water.

Four plankton elements, in the sense of Gran, could tentatively be identified: (1) a group of species, mostly very abundant, found only where the upwelled water was most concentrated; (2) a group, some very abundant, found at almost all stations within the well marked patches of generally high population density; (3) a group of species, none of which reached great abundance but which were relatively more abundant in water which had been on the surface a long time; and (4) a group of species rare throughout the cruise area.

REFERENCES CITED

Allen, W. E.
(In press) Seasonal occurrences and geographical distribution of marine plankton diatoms of southern California in 1941.

Braarud, Trygve
1935. The "Øst" Expedition to the Denmark Strait 1929. II. The phytoplankton and its conditions of growth. Hvalråd. Skr., No. 10: 1–173.

Fleming, R. H.
1939. The control of diatom populations by grazing. J. Cons. int. Explor. Mer, *14* (2): 210–227.

Gran, H. H.
1902. Das Plankton des norwegischen Nordmeeres von biologischen und hydrographischen Gesichtspunkt behandelt. Rep. Norweg. Fish. Invest., *2* (5): 1–218.

Harvey, H. W.
1947. Manganese and the growth of phytoplankton. J. Mar. biol. Ass.U. K., N. S. *26:* 562–579.

Michael, E. L.
1921. Effect of upwelling water upon the organic fertility of the sea in the region of southern California. Proc. 1st Pan. Pacif. Sci. Congr. 1920 (Spec. Publ. Bishop Mus., No. 7 pv. *2*): 555–595.

Moberg, E. G.
1928. The interrelation between diatoms, their chemical environment, and upwelling in the sea, off the coast of southern California. Proc. nat. Acad. Sci., *14:* 511–518.

Sverdrup, H. U. and Staff
1947. Oceanographic observations on the "E. W. Scripps" cruises of 1941. Rec. Observ. Scripps Inst. Oceanogr., *1:* 249–407.

Sverdrup, H. U. and W. E. Allen
1939. Distribution of diatoms in relation to the character of water masses and currents off southern California in 1938. J. Mar. Res., *2:* 131–144.

Sverdrup, H. U. and R. H. Fleming
1941. The waters off the coast of southern California March to July 1937. Bull. Scripps Inst. Oceanogr., *4:* 261–378.

THE CASE FOR STUDYING NORMAL PATTERNS IN FISHERY BIOLOGY

By

LIONEL A. WALFORD
U. S. Fish and Wildlife Service

It seems a reasonable enough idea that a modern government is fulfilling one of its proper functions when it provides adequately for researches about its natural resources. Surely it is no less than a common sense[1] attitude toward our possessions, needing no justification, that we should seek to know all about them—what they are, how we can use them, how much of them we can use, and what affects them. Applied to the sea fisheries, biological research should be comprehensive enough to cover these resources as a whole, extending *all* along the Atlantic, Gulf and Pacific coasts of the United States, as well as Alaska and the island possessions. It should be a continuing program, designed to determine normal patterns of marine ecology, and to provide useful and necessary information for the distant future as well as for the present. It should be flexible enough to take care of anomalies and yet stable enough to resist being completely diverted by them when they occur.

Unfortunately, it is the anomalies, and not the humdrum norms, which attract the most public interest to provide funds for research. Thus, it seems to be generally true that such understanding as we have about our fishery populations results not so much from research pursued systematically to learn normal patterns as it does from investigations of anomalous or otherwise unsatisfactory conditions needing relief.

A species may appear to be in danger of extermination through over-fishing; or fishermen operating two types of gear may dispute as to whether one of them is unduly destructive; a population of shellfish may suddenly vanish without trace; an epidemic may suddenly break out or masses of dead fishes may wash up on the beaches, the stench of their rotting bodies driving tourists away. But whatever the occasion, it is generally the disturbing condition, preferably a disaster, that arouses people's interest. They hope the Government will solve the problem quickly by acting, say, to stop dumping ammunition at sea, or by passing legislation providing a size limit, or by enacting a law to stop all commercial fishing for a species, or to abolish purse seining, or to abate pollution. The Government usually holds hearings over such questions, which usually bring out such diverse and conflicting opinions on the issues that it becomes necessary to gather some pertinent, objectively gathered facts before reaching a decision. Thus the Government starts an investigation of a species, or a fishery, or a particular situation.

---

[1] Webster's New International Dictionary gives a definition of this phrase that is most pertinent to the argument: "Good, sound, ordinary sense; specif., good judgment or prudence in estimating or managing affairs, especially as free from emotional bias or intellectual subtlety . . ."

The first problem ever taken up by the United States Commissioner of Fisheries, in 1871, is an archetypal example of this pattern. In the 1860's people had been fearing that the stocks of fishes in the sea off the coast of New England were diminishing, some of them even to the point of being in danger of extermination. They said that the country was growing, the construction of railroads and the use of ice for packing were making it easier to transport fish long distances; the use of fish for the extraction of oil increased the demand for species like menhaden far beyond anything ever before experienced; and all these things made the danger of depleting the supply all the worse. Some people urged that the simplest way to protect the fish was to abolish traps or pounds, since they were obviously such efficient and unselective gear. Others argued that the evidence against the traps was too flimsy to warrant so drastic a prohibition. Some even asserted there was no sign of diminution at all. This confusing question became the center of local political storms. First, state legislative committees fought over the question without finding an answer to please everyone. Then the United States Congress took it up and in 1871 set up the Commission of Fish and Fisheries to investigate it impartially.[2]

I am going to quote here what Spencer Baird, the first Commissioner, wrote in his first report on this problem, because it shows how clearly he saw the intricate interrelations that he must understand in order to answer the problem intelligently. He said,

> . . . it was a matter worthy of serious inquiry whether so positive a measure as absolute prohibition was expedient or necessary, and whether by limiting the time during which the use of nets is allowed, the interests of both parties may not be reconciled, by giving to the fish the opportunity of spawning undisturbed, and also by regulating the size of the mesh, so as to catch only the oldest and largest fish. All this, however, was only to be ascertained by a careful study of the habits of the fish, so as to determine the nature of their food, the growth of their spawn, and other circumstances bearing upon the solution of the problem in question.
>
> The plan adopted for the inquiry was determined upon after careful deliberation. The great contrariety of opinion developed in the State investigations as to what

---

[2] As passed, the resolution extended the subject of inquiry to the Great Lakes in order to investigate the diminution of white fish and other species.

> should have been the best-known facts in the life-history of the fishes and their associates in the sea, made it necessary to study the natural history of these species as thoroughly as possible, . . .
>
> As the history of the fishes themselves would not be complete without a thorough knowledge of their associates in the sea, especially such as prey upon them or in turn constitute their food, it was considered necessary to prosecute searching inquiries on these points, especially as one supposed cause of the diminution of the fishes was the alleged decrease or displacement of the objects upon which they subsist. . . .
>
> Furthermore, it was thought likely that peculiarities in the temperature of the water at different depths, its chemical constitution, the percentage of carbonic-acid gas and of ordinary air, its currents, etc., might all bear an important part in the general sum of influences upon the fisheries; and the inquiry, therefore, ultimately resolved itself into an investigation of the chemical and physical character of the water, and of the natural history of its inhabitants, whether animal or vegetable. It was considered expedient to omit nothing, however trivial or obscure, that might tend to throw light upon the subject of inquiry, especially as without such exhaustive investigation it would be impossible to determine what were the agencies which exercised the predominant influences upon the economy of the fisheries.

What seems to me to be most interesting about this passage, written three quarters of a century ago, is not quaintness of expression, nor even the general similarity between it and what any fishery researcher might say today in planning his progeam, but rather the measure it gives of the short distance we have progressed in furthering for the country as a whole the desirable aims which Baird implied must be reached to solve such problems. He started with little. There was no great literature about American fishes to guide or hinder him, no systematic meteorological or hydrographic records (except surface temperatures of the sea water in the Gulf of Maine) to associate with past biological observations, even if there had been any; and no dependable catch statistics. That these were desirable, if not necessary data, he recognized at once, and the lack of them must have exasperated him in his efforts to solve his problem objectively. But since he did not have them, he had to get along as best he could, by circulating a questionnaire and otherwise taking testimony from informed or at least opinioned people, and by making collections and surveys of the fauna and of the environment. Within a year, he submitted his recommendations.

This is the pattern which has been repeated in various localities to solve special problems during the whole history of our biological fishery research. There develops an anomalous condition (in this case diminution of the fish stocks below a level which people remembered as having once been much higher); an interested special group of people request that the condition be investigated; after due legislative procedure, scientists are assigned to the problem; to understand the cause of the undesirable condition, the scientists first try to learn about the time when the condition was satisfactory (*i. e.*, the normal pattern), but because systematic past records are nearly always fragmentary or lacking, this effort usually proves fruitless. Then, because they are expected to devise a remedy for the condition in a reasonable time, they make deductions and recommendations from what data they can assemble. Such an investigation may not be conducive to learning much about the normal, because it is bound by too many limitations, for the anomalous condition is usually sharply delimited in scope. It is limited in time to the memory of the current generation, often even to such a short period as a season or two. It is limited ecologically to the affected species which are of most economic value.

The net effect of our preoccupation with problems of this kind is that we neither cover enough ground in our research programs or make fast enough progress toward the ideal goal set forth in the beginning of this paper. Since 1871, we have arrived this far: Of the 70 species of sea fish populations that now yield catches of two million or more pounds a year, we still know virtually nothing about more than half of them. Occasional brief studies have given us a smattering of knowledge about a third of them; and intensive, long range studies have taught us a good deal, though not all the vital things we need to know about the remainder. Furthermore, fishery studies have been very unevenly distributed geographically, so that there are long stretches of coastline which have been relatively neglected. This is too small an accomplishment for 75 years of sea fishery research in a country like the United States. What can we do to speed up the rate of progress? We can be sure that the most dramatic events, the undesirable situations, the anomalies, will continue to generate public support for special investigations and we will have to continue to conduct them. But at the same time, we must try by all means to get better support for the systematic, less spectacular studies of normal conditions, which in the long run will give us more efficiently what we need to know about the anomalous situations.

Meanwhile, there are signs of a reawakening, among biologists, in the appreciation of the cosmic nature of marine fishery problems. For example, the fishery conservation agencies of the states of Washington, Oregon and California, in a recent report to the Pacific Marine Fisheries Commission said,[3] "To collect adequate information for an intelligent management of Pacific Coast fisheries, it will eventually be necessary to set up a program for the collection of data on changes in physical and chemical oceanography. Such environmental factors

[3] Quoted in *Pacific Fisherman*, August 1948: 30.

have direct bearing on the availability of fish on the fishing grounds."

Fot its August 1948 issue, the *Pacific Fisherman* asked several biologists the question, "What does research need?" Some of the replies, anonymously quoted, were as follows:

> Applied research, the study of special problems of fishery biology, needs a background of basic research—the kind that doesn't pay. We lack basic oceanography in the Pacific, where we know little of the upwellings of the deep, nutrient waters that feed the 'feed' which the fish follow, and on which they live. . . .
>
> Research needs adequacy. As is true in many fields of endeavor, efforts of less than a certain magnitude are often ineffectual efforts.
>
> Freedon of action for the trained intelligence. . . .

And another, succinctly expressing the view which has for so long prevailed in this country, said, "Research needs to be concentrated for best accomplishment. The dangerous situations must be attacked with all possible energy. Research should not become diffuse and superficial. World-wide skimming of the surface may allow damage to develop undectected . . ."

This is still a matter of opinion, of course. My view is that the choice is not between handling dangerous situations and being diffuse or superficial. It is rather between studying the normal mechanism of the marine organism as a whole and studying the abnormal of the separate parts, one part after another. In a few years we will have some experience which should take this question out of the realm of dialectics, for several large-scale fishery studies, not tied to anomalous situations or to particular species, are being started. The most spectacular of these is the Pacific oceanic fisheries investigations which the Fish and Wildlife Service is beginning this year (1948) under the terms of the Farrington Act. The purpose of this work will be to explore, investigate and develop the fisheries of the territories and island possessions of the United States in the tropical and subtropical Pacific.

And in another part of the Western Hemisphere, the new program of the recently reactivated Woods Hole Laboratory is devoted to a study of the ecology of the fishery populations of the New England Banks. Here we come around full circle to the same point where Spencer Baird started his work in 1871, and to pretty much the same point of view that he had. Let us now see what progress we can make with this fresh start.

ESTIMATIONS OF ABUNDANCE OF THE EGGS OF THE PACIFIC PILCHARD (*SARDINOPS CAERULEA*) OFF SOUTHERN CALIFORNIA DURING 1940 AND 1941[1]

By

OSCAR E. SETTE AND ELBERT H. AHLSTROM

Fish and Wildlife Service
United States Department of the Interior

INTRODUCTION

One of the more important items of information required for an understanding of the dynamics of population numbers in marine fish stocks is the rate of replenishment of the stock through its reproductive process. It has been established for the Pacific pilchard, or California sardine (*Sardinops caerulea*), that some spawning seasons give rise to much larger year-classes than others (Clark, 1939), as measured at the time the year-class reaches commercial age. But it remains to be determined whether this is due to inequalities in the reproductive process or to variations in the survival rate through any one of the several stages of existence between egg and commercial age.

A requirement for this problem in its entirety is the determination of numbers in each year-class through the several stages: egg, larva, post larva, and juvenile. The first two stages, and perhaps a large portion of the third, are planktonic, and for these stages the problem becomes one in quantitative plankton research. This report is concerned with the enumeration of the eggs in a certain area during two spawning seasons. The eggs, being nonmotile, are perhaps the simplest of the several stages to deal with, yet even their enumeration involves a number of complexities and untested assumptions. The results reported here form a basis for further advances and can in no sense be accepted as final determinations.

The data on which this report is based have been published by Ahlstrom (1948). The hydrographic data for the cruises in which the plankton was collected are given by Sverdrup and Staff (1944, 1947). The sea work was done in collaboration with the Scripps Institution of Oceanography, on its research vessel E. W. SCRIPPS.[2]

---

[1] Published by permission of the Director, U. S. Fish and Wildlife Service.

[2] The authors gratefully acknowledge the aid given by Milner B. Schaefer in the development of the mathematical treatment.

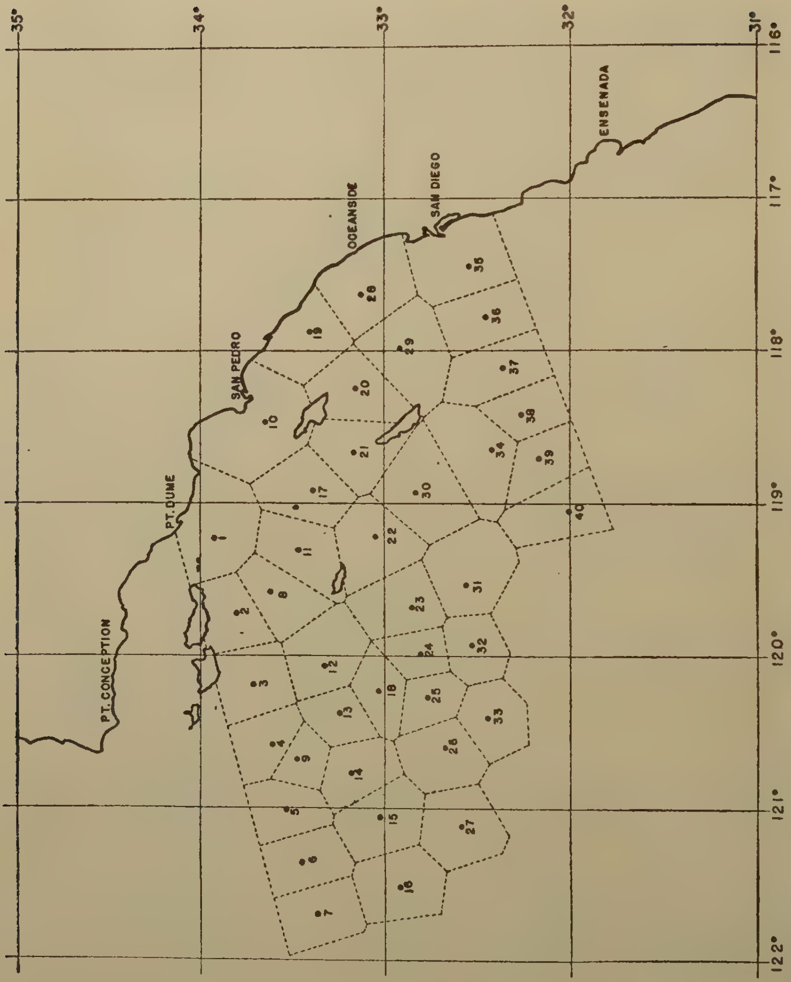

35°
34°
33°
32°
31°
116°
117°
118°
119°
120°
121°
122°
ENSENADA
SAN DIEGO
OCEANSIDE
SAN PEDRO
PT. DUME
PT. CONCEPTION

Figure 1. Chart of cruise XIII, showing the station pattern followed during the 1940 survey and the polygons used for area weighting.

The spawning range of the pilchard, though not definitely known, may extend from southern Lower California to at least as far north as off Central Oregon, and in some places as far seaward as 300 miles. Since this entire range could not be surveyed by the single vessel available, the surveys of 1940 and 1941 were confined to a limited area off southern California which was known to include important amounts of spawning (Scofield, 1934).

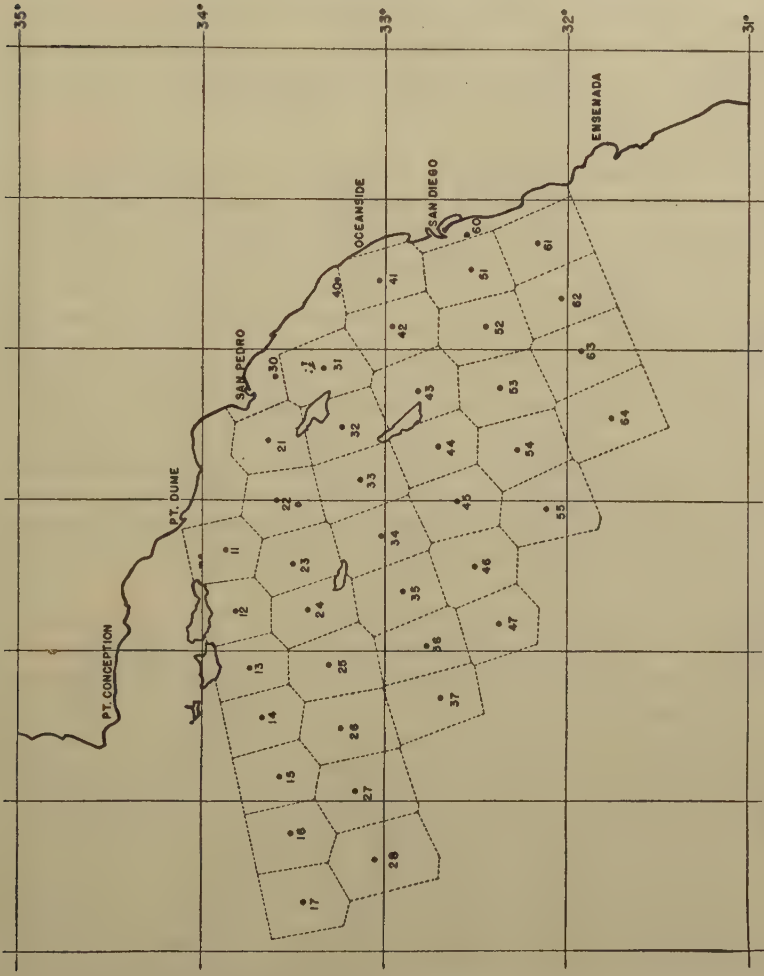

35°
34°
33°
32°
31°
116°
117°
118°
119°
120°
121°
122°
ENSENADA
SAN DIEGO
OCEANSIDE
SAN PEDRO
PT. DUME
PT. CONCEPTION
11
12
13
14
15
16
17
21
22
23
24
25
26
27
28
30
31
32
33
34
35
36
37
40
41
42
43
44
45
46
47
51
52
53
54
55
60
61
62
63
64

Figure 2. Chart showing the station pattern followed during the 1941 survey and the polygons used for area weighting.

A pattern of stations was established in this area, as shown in Figs. 1 and 2, for 1940 and 1941, respectively. There were six cruises each season, giving a mean interval between successive observations at a station of about 17 days during 1940 and about 25 days during 1941.

The plankton nets employed during 1940 were 1.6 meters in diameter at the mouth, while during 1941 they were 1.0 meter in diameter. The nets were constructed of either No. 24xxx grit gauze or of cotton scrim of similar mesh size. Both materials were of fine enough mesh to retain all pilchard eggs. In taking a haul, the net was lowered to approximately 75 meters[3] in 2 to 4 minutes time and then hauled obliquely upward at a slow, uniform speed for approximately 20 minutes. Water flow through a net was measured by a current meter fastened in the center of the mouth.

Because of differences in the speed of the boat during hauling, in the length of time of hauling, and in the degree of clogging of the nets, the amount of water strained varied from haul to haul even for nets of the same size and construction. Also, nets of two different sizes were used during the surveys, the larger straining about two and a half times as much water as the smaller. For comparability, the numbers of eggs in all hauls were converted to the basis of numbers per standard volume of water strained per unit of depth fished. The standard chosen was 10 cubic meters of water strained per meter of depth fished, this particular value being used because it gave a factor of approximately 1 for the 1.0-meter net. In effect, this adjusts the number of eggs in a haul to an equivalent of the number under 10 square meters of sea surface. This area will hereafter be referred to as a "standard area." The procedures used and the resulting standard numbers for each station are listed in Ahlstrom (1948).

Our surveys did not begin early enough in either year to determine with certainty when pilchard spawning commenced off southern California. The first cruise (cruise 10) of the 1940 survey took place beween March 5 and 14. Pilchard eggs were taken in every haul during this cruise, hence spawning was occurring throughout the area and must have commenced some time previously. Some indication of the length of the time involved may be inferred from the data on pilchard larvae. Specimens up to 16.5 mm. in length were collected during cruise 10, suggesting that spawning had been under way for at least a month. The first cruise of the 1941 survey took place two weeks later than in 1940 when the spawning season was even more advanced. The surveys, especially that of 1941, covered the end of the spawning season off southern California more adequately. During the last cruise of 1940, which took place May 27 to June 7, only about 1/20th as many eggs were taken as in midseason. Even fewer eggs were collected on the last two cruises of 1941, which took place after early June, their numbers being only about one-third of one per cent as many as were taken on the first cruise of this season.

In both seasons, pilchard spawning was widespread throughout the area surveyed off southern California, especially at the height of the

---

[3] Pilchard eggs do not ordinarily occur this deep: they are most concentrated in the upper 20 meters and very few are found below the 40 meter level.

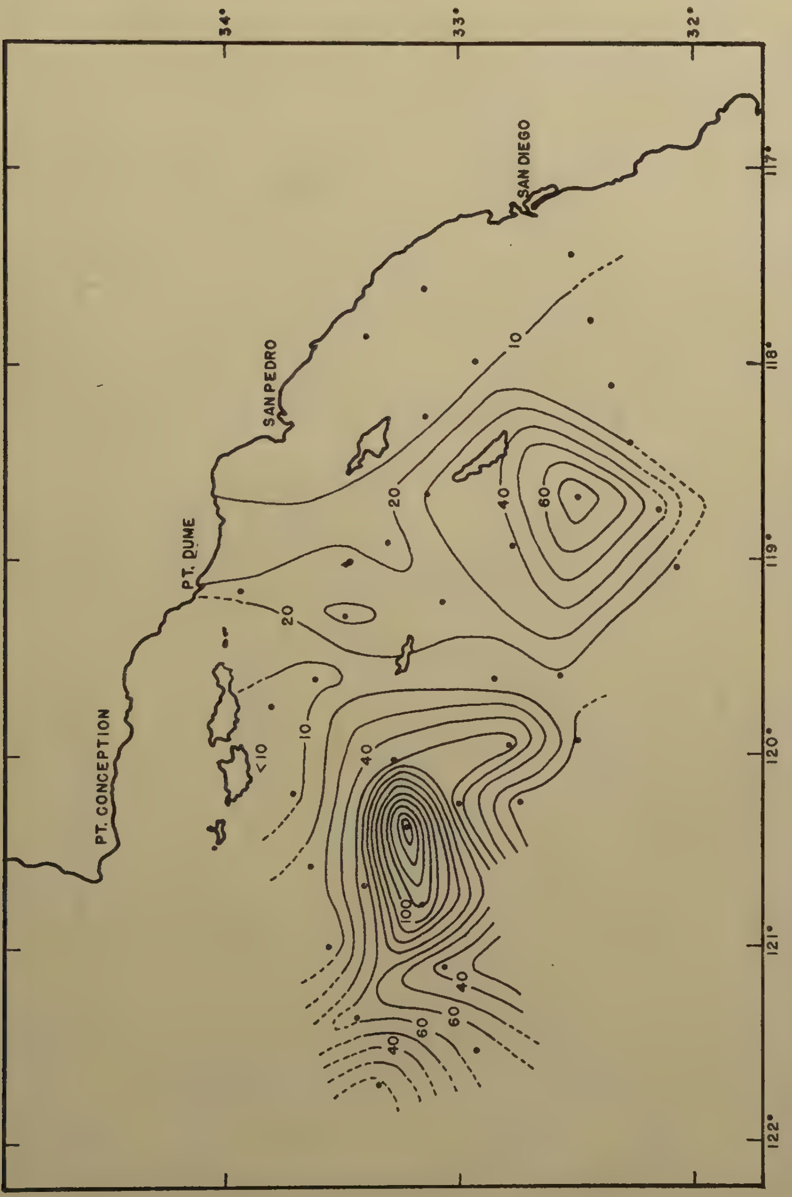

34°
33°
32°
117°
118°
119°
120°
121°
122°
PT. CONCEPTION
PT. DUME
SAN PEDRO
SAN DIEGO
10
20
40
60
100
<10

Figure 3. Chart showing distribution and abundance of pilchard eggs during the period of survey, 1940. The values indicated for isometric lines represent the concentration of eggs, in thousands, per standard area for the period.

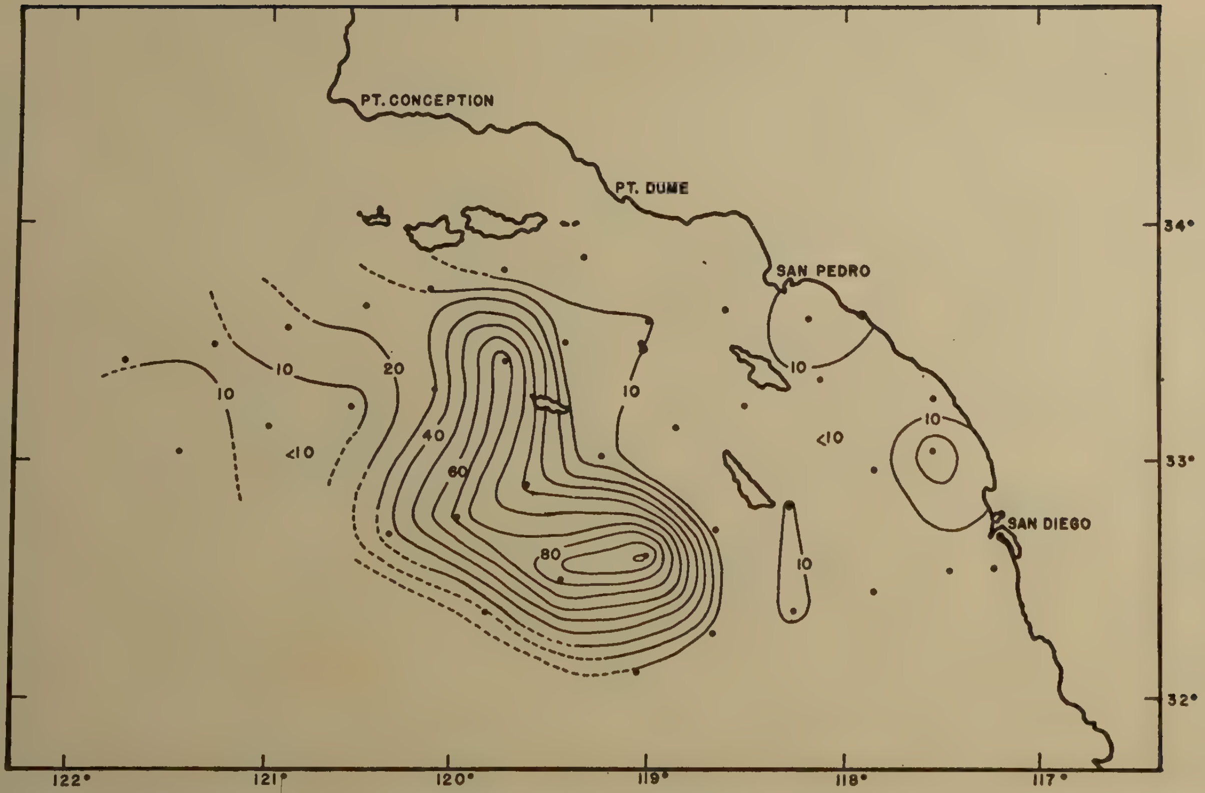

PT. CONCEPTION
PT. DUME
SAN PEDRO
SAN DIEGO
34°
33°
32°
122°
121°
120°
119°
118°
117°
10
10
10
10
10
10
10
<10
<10
20
40
60
80

Figure 4. Chart showing distribution and abundance of pilchard eggs during the period of survey, 1941. The values indicated for isometric lines represent the concentration of eggs, in thousands, per standard area for the period.

spawning season, when pilchard eggs occurred at every station occupied. However, as can be seen in Figs. 3 and 4, spawning was far from uniform throughout the area, most of the spawning occurring just beyond the outlying islands and approximately 60 to 90 miles from the mainland coast. Since the concentration was very high in a relatively small area, there were very high counts of eggs from a relatively few stations. Hence the number of stations that chanced to be included in this area would effect the estimate of total spawning quite materially.

The primary interest of this study is in the total number of eggs spawned rather than in the number of eggs found in the water at any one time. The latter may be high or low depending on the number of days' accumulation of eggs which, given day-after-day spawning, will be greater at lower temperatures when incubation is slower than at higher temperature when it is faster. Extremes of less than 2 to over 4 days were possible, and 2½ to 3½ days were common (Ahlstrom, 1943), in the survey area.

Differences resulting from the effect of temperature on the rate of development can be eliminated by putting estimates of numbers of eggs on a spawned-per-day basis. For one method of computation, a separation was made of the several age-categories[4] present in a sample, and a tabulation was made of the numbers of each.

TREATMENT OF DATA

The basic material available for determining the number of pilchard eggs spawned within the area surveyed each season consists of a set of determinations of egg concentrations composed of six subsets (cruises) distributed about evenly through time, each subset made up of about 40 determinations (stations) distributed about evenly through space. We shall treat the material within each subset as a problem of integration over space and the combining of subsets as an integration on time. In doing so, we will assume that the distribution of egg concentration through space and through time are both continuous, and that the egg-concentration gradients between points in space and time are linear. Even with these simplifying assumptions there are several methods of estimating total numbers for a cruise or for a season.

---

[4] In the original data, Ahlstrom (1948) designated the age categories "A," "B," "C," and "D." In age category "A" were included eggs spawned not longer than 24 hours prior to collection, in "B," eggs spawned within 24 to 48 hours of collection, in "C," eggs spawned within 48 to 72 hours of collection, and in "D," eggs spawned within 72 to 96 hours of collection. Because newly spawned eggs of the current night's spawning are first taken in hauls collected about 8 P. M., this hour was taken to mark the break between age categories. Thus after 8 P. M. all eggs of the previous night's spawning were classed as "B" eggs, all eggs spawned two nights before as "C" eggs, etc.

Definition of Symbols. It will be convenient to adopt a series of symbols to represent various types of estimates and concepts.

c_i: Any of several estimates pertaining to the number of pilchard eggs spawned per day per standard area of 10 square meters of sea surface, at the i-th station.

c_i': The total number of pilchard eggs in a standard haul or standard area at the i-th station.

c_i'': The total number of pilchard eggs in a standard haul at the i-th station belonging to "complete age categories" (as the latter is defined elsewhere), *i. e.*, c' less incomplete age categories.

c_{ia}: The number of eggs in the youngest complete age category at the i-th station.

c_{ib}: The number of eggs in the next older complete age category at the i-th station.

$\bar{c}_i'$: An estimate of c_i, derived by dividing c_i' by d_i.

$\bar{c}_i''$: An estimate of c_i, derived by dividing c_i'' by the number of age categories involved.

c_{it}: The estimated number of eggs spawned in a standard area at the i-th station during the period of survey calculated by

$$c_{it} = D_t \frac{\Sigma c_i'}{h'd''} .$$

C: The estimated number of eggs spawned within the survey area during the period of a cruise calculated by the formula

$$C = D\Sigma\,(w_i c_i) .$$

C': The estimated number of eggs spawned within the survey area during the period of a cruise calculated by

$$C' = 649 \times 10^7 \frac{D}{d'h} \Sigma c_i' ;$$

C_t: The estimated number of eggs in the survey area during the period of survey, *i. e.*, sum of C or C' for all cruises.

C_T: The estimated number of eggs in the survey area during the spawning season.

D: Period of cruise in days.

D_t: Period of survey in days.

d_i: An estimate for each station of the time in days from spawning to hatching, determined from the relation between rate of development and temperature, the latter being the temperature at the 15-meter level at the station.

d': An estimate for the period of a cruise of the average number of days from spawning to hatching, determined from the relation between rate of development and temperature, the latter being the average of temperatures at the 15-meter level for all stations of the cruise.

d'': An estimate for the spawning season of the average number of days from spawning to hatching based on the relation between rate of development and temperature, the latter being the average of temperatures at the 15-meter level for all of the stations included in the season estimate.

h: Number of stations occupied during a cruise.

h': Number of times a station was occupied during the season.

w_i: Any of several weighting factors for the space, in standard areas, represented by the i-th station on a given cruise.

w_i': The weighting factor that is proportional to the area of the polygon assumed to be represented by the i-th station.

w'': The weighting factor based on the assumption of equal weight for each station.

649×10^8: Extent of survey area expressed as square meters of sea surface.

649×10^7: Extent of survey area expressed in standard areas, *i. e.*, units of 10 square meters.

Determination of c_i, the Number of Eggs Spawned per Day. In estimating c_i, the number of pilchard eggs being spawned daily under a standard area (10 square meters) at the i-th station for a given cruise, two methods will be developed, one involving separation of age categories, the other not.

The latter method, being the simpler of the two, will be discussed first. The estimate is obtained by dividing c_i', the total number of eggs in a standardized sample, by d_i, the estimated number of days from spawning to hatching, determined from the relation between rate of development and temperature. Estimates of c_i obtained by this method are designated $\bar{c}_i'$. The temperature of the 10–20-meter level was chosen because most pilchard eggs occur above 20 meters, with the largest concentration between 10 to 20 meters. We have used Ahlstrom's (1943) table 4 for values on the relation between temperature and the rate of development of pilchard eggs.

When basing estimates on individual age categories, one is faced with the problem of incomplete age categories resulting from collection being made while spawning or hatching was actively taking place. Since it can seldom be determined from an examination of the samples whether or not an age category is incomplete due to these causes, the decision has to be arrived at by other means.

From evidence based on the time of collection of newly-spawned eggs (incidence of stage I eggs, see Ahlstrom, 1943), most spawning seems to be confined to the period 8 P. M. to midnight. However, on comparing the number of "A" eggs with "B" eggs in the samples of the 1940 and 1941 seasons grouped by two-hour intervals, we find the following:

Samples collected between:	*Number of samples*	*Per cent that "A" age category is of the "B"*
8–10 P.M.	19	1
10–12 P.M.	23	3
0–2 A.M.	14	9
2–4 A.M.	23	34
4–6 A.M.	17	39
6–8 A.M.	28	125
8–10 A.M.	20	102

From this tabulation it is evident that the "A" age category was caught in markedly smaller numbers than the "B" category between 8 P. M. (when "A" eggs are first taken in hauls) and 2 A. M. During the next four hours "A" eggs, although somewhat more numerous, were still much less abundant than "B" eggs. The reason for this is not clear, since evidence, based on the stage of development reached by "A" eggs collected during this period, indicates that the majority of them must have been spawned several hours earlier. Although we cannot assign a definite reason for the under-representation of "A" eggs in samples collected during this period, it is too marked to be ignored. Consequently, "A" eggs collected between 8 P. M. and 6 A. M. were judged to be incompletely represented in the material, and these were excluded from estimates of c_i'', $\bar{c}_i''$, c_{ia}, and c_{ib}.

When this criterion for incomplete categories is adopted, the classifications of eggs into A, B, C, and D as given by Ahlstrom (1943) must be revised. The former "A" category, collected between 8 P. M. and 6 A. M., is omitted; the "B," "C" and "D" categories taken in the same hauls become, for purposes of this report, "A," "B" and "C" categories, respectively. Those taken at other hours are unaffected.

An age category may also be incomplete due to a portion of the eggs having hatched. For distinguishing these, the assumption was made that if 80% or more of the eggs of an age category were in the stage immediately preceding hatching (stage XI; for a description see Ahlstrom, 1943), a portion of the eggs of the category would have hatched already; the category would be judged incomplete, and such groups of "C" or "D" age categories would be excluded from estimates of c_i'', $\bar{c}_i''$, c_{ia}, and c_{ib}.

Having eliminated incomplete age categories, an estimate may be based on "A" eggs alone, "B" eggs alone, or the mean of all complete age categories of eggs. These are designated as c_{ia}, c_{ib}, and c_i'', respectively. No separate estimates were made for older age categories because, after elimination of incomplete age categories, there were only the two categories "A" and "B" left at many stations where the temperature was above 15° C, and less than three days were necessary for incubation.

Determination of w_j, *the Areas of Stations.* The concept for "area of station" is a convenience for one method of integration over space. It involves the idea that catch at a particular station is not only an estimate of egg concentration at that particular point but also an estimate of concentration over some area surrounding the particular point. Buchanan-Wollaston (1937) obviously was influenced by this idea in suggesting that a relative estimate of abundance of plankton in sets of hauls be made by weighting the catch at each station by a number proportional to the area of a circle whose radius was equal to the mean distance between the given station and adjacent ones. This method has two disadvantages: It is not always possible to define "adjacent station" satisfactorily, and the sum of the weighted catches is relative rather than absolute because the sum of the areas of the circles will not equal the area of the portion of the sea over which the stations were distributed.

However, if instead of circles a system of polygons be adopted, in which the polygon enclosing each station is erected by constructing the perpendicular bisectors of lines drawn from the station to each of all surrounding stations, it will be found that there is a unique polygon surrounding each station (except the peripheral ones). If the concentration at the station then be assumed to extend to the limits set off by each polygon, a system of weightings proportional to the areas of the polygons will result in a set of values which, when summed, will give results identical to those which would be obtained by linear interpolation between stations.

Such a system has the advantage that the sum of the weights equals the area surveyed; furthermore, it is completely objective, excepting possibly the treatment of peripheral stations, and as will be shown below, gives results essentially comparable to those from the conventional but less objective method of constructing isometric lines. Moreover, the polygons are easily constructed when one finds the perpendicular bisectors by swinging arcs of equal radii, the latter selected to be slightly greater than one half the distance between a given station and its neighbor. The areas of the polygons may be determined mathematically or by planimeter. With the latter, the work is about equal to that required when measuring the areas bounded by lines in an isometric system.

For these various reasons the above described system of polygons was adopted for our estimates involving area weightings, and in these the weight, w_i', is the area of the polygon corresponding to the i-th station on a given cruise in 10-square-meter units, *i. e.*, "standard areas" as earlier defined. For peripheral stations, where coast lines were not involved, the outer boundary of the northern line was placed north from this line at half the average distance from this line to the next adjacent line to the south; the outer boundary of the southern line was placed south of the line at half the average distance from the line to the next adjacent line to the north; a station located at the end of a line was enclosed by assuming a station to be present on the line beyond the end station at a distance equal to the average distance separating stations on the line. The area of the resulting polygons was measured by a planimeter, using the coastline as a boundary where it overlapped a polygon.

Because one or more stations were omitted in some of the cruises, polygons were constructed for each cruise of the 1940 survey. An example of a complete one is given in Fig. 1. In 1941 no stations were omitted and a single system of polygons was erected for that season's station pattern, using the average position of occupying each station, as shown in Fig. 2.

It should be noted that the area surveyed during 1941 differed in several respects from the area surveyed the preceding season. The chief differences were: (1) Three stations (Nos. 26, 27, and 33) in the southwest portion of the 1940 survey pattern had no counterparts in the 1941 pattern, and (2) the southern line of four stations (Nos. 61 to 64) of the 1941 pattern had no counterpart in 1940. These stations without counterparts were omitted. The 1941 survey area then amounted to approximately 18,900 nautical square miles. The area surveyed during each cruise of the 1940 season was as follows:

Cruise Number	*Area (nautical sq. mi.)*	*Factor to adjust area to 18,900 nautical sq. mi.*
10	17,450	1.083
11	17,100	1.107
12	19,000	0.997
13	19,150	0.987
14	19,800	0.956
15	19,700	0.960

For comparison with each other and with the 1941 cruises, the 1940 cruises were adjusted to an area of 18,900 nautical square miles, using the above-listed factors. In so adjusting these cruises, we are assuming that the average abundance in a nonsurveyed portion was equal to the average abundances in a surveyed portion.

A more rapid, although less exact, estimate of egg numbers in the survey area may be made by weighting each station equally. Where this has been done, the symbol w'' is used. For comparability of the results obtained by this method with others, the total was put on the basis of 18,900 nautical square miles (649×10^8 square meters) and stations without counterparts in both years were excluded. For a given cruise, the value of w'' was determined by dividing 649×10^7, the number of standard areas of 10 square meters each in the survey area, by the number of stations occupied.

Determination of D, the Time Period of a Cruise. In integrating egg numbers over time, the same concepts were used as in integration over space; that is, the egg concentration observed at a particular time was assumed to represent an estimate for an adjacent period of time also. Specifically, each cruise was assumed to represent a certain period, the exact number of days in this period, D, being the sum of the number of days spent in making the cruise in question plus one half of the days intervening between the end of the preceding cruise and the start of the cruise under consideration, plus one half the time interval in days between the end of the cruise under consideration and the start of the following cruise. The time interval preceding the first cruise of a season was assumed to represent half as long a time as the interval between the first and second cruises. Similarly, the time interval following the last cruise of a season was assumed to represent half as long a time as the time interval between it and the preceding cruise.

By this method of determining D, the result is a computed total which is equivalent to one employing a linear interpolation in time, just as was true with respect to space for the determinations of w_i.

Estimates of Total Eggs Spawned. The foregoing sections give four methods of determining c_i and two for determining w_i; hence, there are eight possible estimates available for each cruise. By adding the appropriate estimates for the six cruises for each season, estimates of C_t, the number of pilchard eggs in the survey area during the period of the survey have been obtained; these are given in Table I. However, because spawning had commenced in the survey area some time prior to the start of the first cruise of each season, the estimates of pilchard eggs in the area during the period of the survey did not represent the number of eggs spawned during the complete spawning season in either year. From the number and sizes of larvae taken during the first cruise of each season, it was estimated, by a procedure to be reported upon in a subsequent paper, that 11% in 1940 and 17% in 1941 of the spawning had taken place prior to the period covered by the survey estimate. The season totals after adding these percentages are denoted by C_T.

Simplified Procedure for Estimating the Amount of Spawning in the Survey Area Based on Average Conditions During a Cruise. The methods previously described require the separate determination of c_i and w_i' for each station. A computation that does not require this is:

$$C' = 649 \times 10^7 \frac{D}{d'h} \Sigma c_i' ,$$

where use is made of the sum of the standard haul totals for all stations occupied during the cruise, put on a per-day basis by dividing this value by the number of days required for the eggs to develop from spawning to hatching at the average temperature of the 15-meter level during the cruise, and C_t by summing the C' for all cruises of the season.

Determination using Isometric Lines. In practice, when using the conventional method based on isometric lines, the number of lines must be kept relatively few, and there is some subjectivity in drawing the lines, hence the pattern follows linear interpolation approximately rather than precisely. Nonetheless, for showing distribution, the isometric chart is nearly indispensable, and it can also serve for estimating total quantities. Having a single chart showing the abundance during each year, our estimation was done for the season as a whole rather than by individual cruises. The values for each station, given in Figs. 3 and 4, represent estimates of the number of eggs spawned during the period of survey in a standard area, *i. e.*, c_t as defined previously. The area between lines was determined by planimeter. The estimates obtained by this method and given in the last column of Table I are similar to those obtained by the other methods.

TABLE I. COMPARISON OF ESTIMATES OF NUMBER OF PILCHARD EGGS, IN BILLIONS, SPAWNED IN AREA SURVEYED OFF SOUTHERN CALIFORNIA DURING SPAWNING SEASONS OF 1940 AND 1941

Cruise No. and Date	*Estimates Based on Formulation $C = D\Sigma(w_i c_i)$ Using*								C'	*Average*	*Using Isometric Lines*
	$w_i' c_{ia}$	$w_i' c_{ib}$	$w'' c_{ia}'$	$w'' c_{ib}$	$w_i' \bar{c}_i''$	$w'' \bar{c}_i''$	$w_i' \bar{c}_i'$	$w'' \bar{c}_i'$			
1940:											
10 (III:5–14)	29869	28224	33082	29590	27459	30220	25848	27626	27843	28862	
11 (III:20–28)	22540	32566	23074	26588	31144	28614	30765	28431	29073	28088	
12 (IV:4–14)	22231	15722	25256	17182	23457	24704	22567	23367	23311	21977	
13 (IV:22–V:3)	51830	54046	62716	61900	52452	62007	48299	58145	57933	56592	
14 (V:10–21)	29611	30314	30826	35290	31779	34377	30260	33009	35257	32303	
15 (V:27–VI:7)	3126	2321	3674	2745	2832	3240	2410	2702	2717	2863	
Survey total, C_t	159207	163193	178628	173295	169123	183162	160149	173280	176134	170685	180929
Season total, C_T	178884	183363	200706	194713	190026	205800	179943	194697	197903	191781	203291
1941:											
17 (III:21–30)	97185	61531	95282	62432	78515	77368	68024	66983	69009	75148	
19 (IV:14–25)	14932	67001	13542	61549	39649	36452	38509	35979	38350	38440	
21 (V:8–17)	10584	16642	11464	17245	12801	13385	13625	14206	15230	13909	
23 (VI:2–12)	5510	4053	5442	3716	4601	4425	4669	4474	4582	4608	
25 (VI:25–VII:3)	442	103	409	110	268	256	254	244	242	259	
27 (VII:21–28)	33	5	36	10	28	22	23	28	21	23	
Survey total, C_t	128686	149335	126175	145062	135862	131908	125104	121914	127434	132387	128335
Season total, C_T	155043	179922	152018	174773	163689	158925	150728	146884	153535	159502	154620

Comparison of Estimates. The estimates of the number of pilchard eggs spawned in the survey area during the spawning seasons of 1940 and 1941 are tabulated in Table I. The lowest estimate for the 1940 season is 87% of the highest and 93% of the average estimate; the lowest estimate for the 1941 season is 82% of the highest and 92% of the average estimate.

For each season, a group of four estimates differs from a second group only in the method of determining w. The estimates based on w' attempt to weight each station according to the portion of the survey area represented by it; those based on w'' give equal weight to the stations. A comparison of these estimates, in billions of eggs, follows:

	1940 Season			*1941 Season*		
Value of	*Estimate using*			*Estimate using*		
c used	w'	w''	w'/w''	w'	w''	w'/w''
c_a	178,884	200,706	.891	155,043	152,018	1.020
c_b	183,363	194,713	.942	179,922	174,773	1.029
$\bar{c}''$	190,026	205,800	.924	163,689	158,925	1.030
$\bar{c}'$	179,943	194,697	.924	150,728	146,884	1.026
Average	183,054	198,979	.920	162,345	158,150	1.027

The two groups of estimates agree more closely for the 1941 season than for the 1940 season. This is not surprising in view of the fact that the 1940 station pattern was more irregular than the 1941 pattern.

A comparison of estimates, in billions of eggs, using different methods of determining c, shows the following:

1940

	c_a	c_b	$\bar{c}''$	$\bar{c}'$	*average*
w'	178,884	183,363	190,026	179,943	183,054
w''	200,706	194,713	205,800	194,697	198,979
Average	189,795	189,038	197,913	187,320	191,016

1941

	c_a	c_b	$\bar{c}''$	$\bar{c}'$	*average*
w'	155,043	179,922	163,689	150,728	162,345
w''	152,018	174,773	158,925	146,884	158,150
Average	153,530	177,347	161,307	148,806	160,248

All estimates of c agree quite well for the 1940 season, less well for the 1941 season. The poorer agreement for the latter season may be partly the result of fewer samples having been taken during the period of more intensive spawning during this season. The $\bar{c}'$ estimates are somewhat lower than those based on age categories: 3% for 1940, 9% for 1941.

The simplified procedure, C', which can be calculated much more rapidly than any other given in the paper, is within 4% of the average estimate for each season, and would alone suffice as a determination of the number of eggs spawned in the area during these seasons.

RELIABILITY OF THE ESTIMATES

A number of counts and measurements enter into the process of computing the total number of eggs spawned in the survey area. Each of the counts or measurements represents more or less imperfectly the true value of the measurement sought. In other words, each contributes to the variability or statistical error of the final total. The two principal mensuration categories are 1) the concentration of eggs at points in time and space, and 2) the locations of the points in time and space. In addition to these there are the approximations involved in interpolations.

Variability of Egg-concentration Data. Sampling of the survey area was done at points (stations) that were about evenly distributed rather than randomly distributed through space and time. Likewise, the egg concentrations that were being sampled were not randomly distributed. The nonrandomness of the distribution in space may be seen by referring to Figs. 3 and 4, and in time by comparing the cruise totals given in Table I. The concentration of eggs at a given point was measured by making a plankton haul, sorting and counting subsamples of the catch, and subsequently converting the count to represent the number of eggs in a standard area.

Due to the nonrandom nature of the distribution of measurement points and of the eggs themselves, the statistical error of individual hauls cannot be measured by root mean sums of squares of deviations from the grand mean. Instead, each datum on egg concentration must be considered as a sample drawn from a different universe, or alternatively, the entire collection of egg concentration might be regarded as samples stratified in time and space. However, the latter alternative is not a promising approach to the study of reliability unless values both of the egg concentrations and of their variances can be expressed as functions of time and space (or in relation to some other frame of reference such as possibly physical, chemical or biological qualities of the sea water in which the eggs were taken and which might determine their concentration through their influence on the act of spawning). Accordingly we shall examine the data from the standpoint suggested in the first alternative, that is, as if each concentration datum had been drawn from its own peculiar population.

Fortunately, the possibility of the need for this study was foreseen in planning the surveys, when "paired hauls" were included in the design of observation for five of the 1940 cruises. By the method of analysis of variance of egg counts in a group of the paired hauls of 1940, selected for containing more than a certain minimum number of eggs, Silliman (1946) already has studied the variances ascribable to differences between stations (time-place of hauls), between members of a paired haul at the same station and between samples counted from the same haul, as well as "haul-order," "sample-order" and interactions. Several features of his findings are of interest here.

First, the variability (mean square) between stations was about twelve times as high as the variability between hauls, proving statistically the point discussed above, that hauls taken at different stations cannot be considered as samples drawn from the same population.

Second, due to the markedly skewed distribution of numbers of eggs per sample, Silliman found it necessary to use the logarithms of the egg counts, which he found were nearly normally distributed. In this use of logarithms, Silliman followed Winsor and Walford (1936), and Winsor and Clarke (1940). The latter, in considering the counts of the various plankton organisms in each of twelve oblique hauls made in succession and considered as a sample of the same body of water, had found that the range of variation was very different for the different groups of plankton organisms, the range being roughly proportional to the mean for each group. They also observed a tendency for the numbers of different kinds of organisms in any haul to be high or low together, indicating factors common to the whole haul affecting all groups in the same direction. They concluded that "under these circumstances, it is both convenient and theoretically advantageous to use the logarithm of the catches rather than the catches themselves" (Winsor and Clarke, 1940: 4).

Since Silliman carried through his analysis in logarithms, his estimate, say of 0.024 for variance between hauls, is the variance of the logarithms of egg-count, not of the counts themselves. Similarly, the square root of the variance, 0.155, is the standard deviation of the logarithms of egg-counts. The antilog of this standard deviation, 1.429, stands in relation to the egg-counts themselves as a ratio by which the egg-count is multiplied or divided to get the upper and lower one-standard-deviation limits, respectively, of the sampling variability of the egg-count.

These upper and lower limits are not equidistant from the actual value of the egg-count. If we let the logarithm of egg-count equal zero, its antilog will equal one, and the upper limit is .429 while the lower is $1 - 1/1.429 = .300$ from the value (one) of the egg-count. This asymmetry, of course, is due to the fact that when the distribution of the logarithms of a variable is normal, the distribution of the variable itself is skewed, the curve falling more sharply to the left and more gradually to the right of the mode. Attending this asymmetry, of course, is the difference between the antilog of the mean of the distribution of the logarithms of the variable and the mean of the distribution of the variable itself. In this case the antilog of the mean of the logarithms of the egg-counts is one, as stated. The mean of the actual egg-counts themselves is approximately 1.06. This will be recognized as a specific example of the familiar fact that the geometric mean (antilog of the mean of logarithms of numbers) is always somewhat lower than the arithmetic mean of a series of numbers. It is of interest here, because it suggests that if the sampling variation of numbers of eggs in hauls is such that their logarithms are distributed normally, the mean of these numbers will be somewhat higher than its most probable or modal value.

Of greater significance is the fact that the sampling variation of numbers of eggs in hauls is proportional to the number in the haul. In the case of one egg in the haul, we have seen that the upper and lower one-standard-deviation limits are .429 and .300, respectively, from the value one. For 100 eggs the corresponding limits are 42.9 and 30.0, respectively. The one-standard-deviation range is exactly one hundred times as great.

Our estimates of the number of eggs spawned into the survey area are essentially sums of numbers, each of which has the sampling variability of the type discussed above. As indicated by the example given, the sampling variability is large, and as expressed by the standard deviation is proportional to the number itself. Given this kind and degree of variability in the numbers summed, what is the variability of the sum itself?

Silliman (1946: 82), in giving a table of fiduciary limits for means of egg-counts, evidently had this question in mind. But this method of computing fiduciary limits, using logarithms, gives a result that is valid only for geometric means, or for arithmetic means in the special case where the number in the egg-counts is identical for all items entering the mean. It would be a useful approximation for expressing the sampling variation of the mean of a number of hauls made at nearly the same time in nearly the same locality where it might be expected that the number of eggs per haul would be nearly uniform among the several hauls.

With the present case, where the numbers entering the sums range through several orders of magnitude, and with the sampling variability proportional to the number itself, it is obvious that the larger numbers contribute vastly more to the variability of the sums than the smaller numbers, and this must be reflected in the sampling variation of the sums of such numbers.

It is proposed, therefore, to enploy the fact that the variance of a sum of independent variates is equal to the sum of the variances of the variates, and to estimate the variance of a variate from the variance of the logarithms of that variate. In our case the variate is the number of eggs per haul at a station, and we will use the values for the variance of the logarithm of the number of eggs per haul computed by Silliman (1946) for a group of 24 paired hauls of the 1940 survey. Since this group is not necessarily a random sample of the hauls comprising our material, the use of Silliman's results gives only an approximation of the true variances. To be discussed later are the several respects in which the nonrandom nature of the 24 hauls selected by Silliman may influence our estimate.

In estimating the variance of the number of eggs in a haul, it is convenient to take advantage of the fact that it is the standard deviation and not the variance (or standard deviation squared) that is proportional to the number of eggs in the haul. Since the sampling distribution of the logarithms of the numbers is normally distributed, the root-mean-square of the deviation from the mean of the sampling distribution of the corresponding numbers themselves will be an estimate of their standard deviation. This distribution is skewed, of course, and the individual standard deviations may not be given the same interpretation as when computed for a normal distribution. However, when a number of sets of samples are drawn from a population, the distribution of the means, and therefore the sums of the sets, approaches normality even though the population from which drawn be markedly skewed. It will be assumed that this holds true for our sums of egg-counts, where the count itself is analogous to the individual in a sample and where the sampling distribution of hauls is analogous to the population from which the sample is drawn; our problem, then, is to find the standard deviation of the distribution of x where the distribution of log x, denoted by Lx, is normal with the estimated standard deviation of s_{Lx}. A general mathematical solution would involve extensive analysis, but, having to deal with only several values of s_{Lx}, it is fairly simple to erect a curve of distribution of x which corresponds to the curve of distribution of Lx for a given s_{Lx} through use of tabular values for the normal frequency function such as the one given in Hodgman (1944: 200–204). Given this distribution of x, its estimated standard deviation, s_x, may be readily computed by the root-mean-square method.

This was done for haul counts, the mathematical expectation of the logarithm of which is zero (one egg per haul) and values of s_{Lx} appropriate (according to Silliman, 1946) to cases where (1) the egg number in our data was estimated from one haul from which one sample was counted ($s_{Lx} = 0.176$) and (2) where it was estimated from two hauls from which two samples were counted ($s_{Lx} = 0.1183$). Also, it was done for two arbitrary values: $s_{Lx} = 0.10$ and $s_{Lx} = 0.25$. The results are as follows:

s_{Lx}	*Antilog. of* s_{Lx}	s_x	$\bar{x}$	$s_x/\bar{x}$
0.1760	1.500	0.4505	1.0856	0.415
0.1183	1.313	0.274	1.038	0.264
0.1000	1.259	0.224	1.027	0.218
0.2500	1.778	0.732	1.180	0.620

The arbitrary values of s_{Lx} were included to explore the possibility of a linear relationship between s_{Lx} and s_x or associated values. The relation between antilog s_{Lx} and s_x seems to approach linearity and is approximately defined by

$$s_x = .944\,(\text{antilog } s_{Lx}) - .965\,. \tag{1}$$

This formula, calculated from the co-ordinates antilog $s_{Lx} = 1.5$, $s_x = 0.4505$ and antilog $s_{Lx} = 1.313$, $s_x = 0.274$, gives the value $s_x = .227$ for $s_{Lx} = .1000$ and $s_x = .729$ for $s_{Lx} = .250$ as compared with the above-listed values for s_x obtained directly by the root-mean-square calculation. The agreement is as close as would be expected from consideration of computational accuracy, and linear interpolation for other values of s_x may be confidently used within the range $s_{Lx} = 0.1$ and $s_{Lx} = 0.25$.

In estimating the total numbers of eggs spawned into the survey area, the counts of eggs were multiplied by certain factors before getting their sums. Since the standard deviation of the number caught in a haul, $s_{\bar{x}}$, is proportional to $\bar{x}$, the number caught in a haul, as earlier explained,

$$s_{\bar{x}} = s_x \cdot \bar{x}\,,$$

where the mathematical expectation of x equals one, then

$$s_{a\bar{x}} = a \cdot \bar{x}s_x\,,$$

where a represents any weighting factor.

Furthermore, where such weightings have been applied to summed sets of egg-counts, such as weighting the egg numbers for a cruise according to the time-period represented, the standard deviation of the sum, s_s, is determined from the standard deviations of the individual haul counts, $s_{\bar{x}_i}$, by the equation:

$$s_s = (s_{\bar{x}_1}{}^2 + s_{\bar{x}_2}{}^2 + \ldots + s_{\bar{x}_n}{}^2)^{\frac{1}{2}},$$

and, if the counts be multiplied by weighting factors w_i,

$$s_{ws} = (w_1{}^2 s_{\bar{x}_1}{}^2 + w_2{}^2 s_{\bar{x}_2}{}^2 + \ldots w_n s_{\bar{x}_n}{}^2)^{\frac{1}{2}};$$

also, where the w_i are all equal,

$$s_{ws} = w(s_{\bar{x}_1}{}^2 + s_{\bar{x}_2}{}^2 + \ldots + s_{\bar{x}_n}{}^2)^{\frac{1}{2}}.$$

Inasmuch as all of the estimated totals of numbers of eggs spawned in the survey area are based on the same original hauls, the portion of their variability attributable to sampling variations of hauls must be similar, unless the weightings employed were markedly different. It is probably sufficient, therefore, to estimate the haul variability associated with the estimate of C_t in the column headed $w_i'c_i'$ in Table I.

We shall let the several values for $w_i'\bar{c}_i'$ equal x_1, x_2, etc. There were three types of haul-sample combinations in the 1940 and 1941 survey data for which the following values were used:

Station value based on:	*Series Designation*	V_{Lx} *(from Silliman)*	s_{Lx}	*Antilog.* s_{Lx}	s_s *(from Eq. 1)*
1 haul, 2 samples	x'	.0275	.1658	1.465	0.418
1 haul, total count	x''	.024	.1549	1.429	0.384
2 hauls, 4 samples	x'''	.014	.1183	1.313	0.274

The computation for the estimate of total eggs spawned for one season's survey, C_t, was computed by the formula

$$C_t = D_1\Sigma x_i + D_2\Sigma x_i + \ldots + D_n\Sigma x_i,$$

where D is a weight proportional to the number of days' spawnings represented by a cruise and its subscripts the first, second, etc., cruises of the season.

The computation of the standard deviation for sampling variation for the sum of eggs of each cruise, s_C, then is:

$$s_C = (s_{x_1}{}'^2 + s_{x_2}{}'^2 + \ldots + s_{x_n}{}'^2 + s_{x_1}{}''^2 + s_{x_2}{}''^2 + \ldots + s_{x_n}{}''^2 + s_{x_1}{}'''^2 + s_{x_2}{}'''^2 + \ldots + s_{x_n}{}'''^2)^{\frac{1}{2}};$$

and the computation for s_{C_t}, the standard deviation for sampling variation for the sum of the cruises for the season then is:

$$s_{C_t} = (D_1^2 s_{C_1}^2 + D_2^2 s_{C_2}^2 + \ldots + D_n^2 s_{C_n}^2)^{\frac{1}{2}},$$

where the subscripts of D and C denote the first, second, etc., cruises of the season.

The resulting values for C_t and s_{C_t} are given in Table II in two parts: one part consisting of the contributions from stations whose $w'\bar{c}'$ is more than 20 and the other part whose $w'\bar{c}'$ is less than 20, designated in the table as high-concentration stations and low-concentration stations, respectively. The high-concentration stations have egg-counts lying within the range studied by Silliman and the low-concentration stations have counts lying below this range.

It was mentioned earlier that Silliman's estimate of variance, V_{Lx}, was on a nonrandom sample of our hauls, and the most important departure from randomness was his elimination of samples with low egg-counts. In the latter, both the sampling variability of hauls and of samples drawn from the catches made by the hauls for the counting of eggs must be larger relative to the count itself than for those with large counts. Therefore, the contribution to the standard deviation of the sum for all stations is underestimated for the small hauls. However, for both 1940 and 1941 this contribution is so small that even a 100% change in the standard deviation contributed by the low-concentration stations would change the standard deviation of the survey area sum by considerably less than 1%.

There are other respects in which the statistics for V, as given by Silliman, may not estimate correctly the variability of our data. Undoubtedly s_x varies directly with x because there are a number of factors operating on a haul and on the measurements pertaining to the haul *as a whole.* For our particular hauls there are several such factors that can be recognized readily. For instance, Silliman found "haul-order" and "sample-order" variances that could be traced to faults in the mechanics of handling the material caught by the hauls. The discrepancies were such as to tend to be cancelled upon pooling the two members of paired hauls in 1940; the mechanical faults themselves were rectified in 1941 procedures. Hence, these do not materially affect the s_x values employed herein. They serve as examples of factors, unrelated to the actual concentrations of eggs in sea, which can affect the number of eggs supposed to represent the concentration in the sea at a certain time and place.

Similar mechanical faults may affect the hauling process itself. In our hauls the most likely fault is the variation in the amount of water strained at the levels of maximum concentration of eggs. It will be recalled that the hauls were oblique and that the level at which the net was towing at any particular time was controlled by the length of towline and the angle of stray, as reported by Ahlstrom (1948). The length of towline could be controlled precisely enough, but the angle of stray was a function of speed of the ship relative to the water mass acting upon the net and could be controlled only very approximately. Operational experience has shown that sometimes the net towed more than twice as much at one level than at another through the series of some 20 levels that made up each oblique haul. Records of vertical distribution of eggs (Ahlstrom, 1948) also proved that eggs were sharply stratified vertically, with by far the highest concentrations occurring within the upper 20 meters. However, the townet was drawn from about 70 meters to the surface and only a part of the tow took place in the upper 20 meters. While the total amount of water strained during the haul was registered by current meter and computed from initial and final readings, this could reveal nothing as to the amount of water strained in the upper egg-rich layers. This no doubt explains Silliman's (1946: 81–82) failure to find any reduction of variability between paired hauls after both were "standardized" by taking into account the measurements of water passing through the net in each haul.[5] This, we believe, is the principal reason that the sampling variability in our hauls is large and is proportional to the number of eggs in the haul instead of being in accordance with the Poisson series, as might be expected *a priori* and according to which its standard deviation would be proportional to the square root of the number.

Due to improvement in 1941, both in equipment and in procedure, we believe that the sampling variation attributable to the making of hauls was substantially reduced in 1941 as compared with 1940, and in using the values for V based on 1940 material we tend toward placing a maximal value on the standard deviation for the 1941 estimation.

Also, the sampling variation attributable to the drawing of the samples for counting the eggs was lowered in 1941 as compared with 1940, both by using a better fractioning device and by drawing a larger sample. Accordingly, the standard deviation is overestimated as concerns this factor also.

It is our opinion that the three principal respects in which our

[5] The type of haul employed in 1940 and 1941 was designed to reach the maximum depth at which larvae as well as eggs are found. For taking only eggs to reduce sampling variability of hauls, it would have been better to haul the nets from about 20 meters to the surface. However, this would have caused a systematic underrepresentation for eggs and even more serious faults with respect to larvae.

TABLE II. VARIABILITY OF ESTIMATE OF TOTAL EGGS SPAWNED IN SURVEY AREA

Category	*No. of Sta.* (h)	*Est. of eggs spawned* (C_t)	*Stand. dev. of est.* (s_{C_t})	*Coefficient of variation* ($100s_{C_t}/C_t$)
1940:				
High-concentration stations	83	150,534	7,539	5.01
Low-concentration stations	89	9,619	467	4.85
All stations	172*	160,153	7,553	4.71
1941:				
High-concentration stations	45	117,437	12,056	10.27
Low-concentration stations	73	7,656	557	7.28
All stations	118*	125,093	12,069	9.64

* Includes all stations at which pilchard eggs were taken, excludes stations with 0 concentration of eggs.

material is not typical of the material from which our value for V was taken have the net effect of somewhat underestimating the standard deviation of the estimate for 1940 and overestimating it for 1941.

One may conclude, therefore, that a true value for the standard deviation for 1940 would be somewhat over 7,553, and for 1941 somewhat under 12,069, perhaps not over 10,000 billion in either year. The difference of the survey area estimates of the two years is approximately 35,000 billion eggs, and hence the estimates for the two seasons are somewhat over two standard deviations apart. Thus, as far as sampling variability of hauls is concerned, the observed difference between the two seasons is demonstratively significant.

Dimensions in Time and Space. The dimensions in time and space which entered into the estimates as weighting factors were measured with sufficient accuracy, in a physical sense, to be negligible in relation to other sources of variability. However, time enters into the computation fundamentally through its use to measure the duration of the incubation period of the eggs, in the estimates based on $\bar{c}'$, which, it will be recalled, is the catch of eggs divided by the time in days from spawning to hatching and which varies with temperature. The accuracy of such estimates depends, then, on the reliability of the determination of the relation between temperature and development and on the faithfulness with which the temperatures observed at the collecting point reflect the temperature during the entire period of development up to the collecting time. There is opportunity for systematic error in both of these mensuration elements. For this particular set of estimates, any such systematic error would have very little effect on the result, because the scale of development rate was determined from the very same set of data (including the measurement of temperature) as was used for our estimates. In effect, this adjusts the mensuration instrument for observational error, and the relative numbers of eggs estimated to have been spawned should not be affected except for the variability associated with predicting the incubation time from the regression line of time to reach stage XI on temperature at the 10 to 20 meter stratum, as given by Ahlstrom (1943: table 5). Since the standard error of such predictions, given in terms of the reciprocal of the logarithm of the temperatures, is equivalent to approximately 7% of the incubation period, and since some hundreds of such estimates of $\bar{c}'$ are involved in the total survey estimate, this error is negligible. It may be noted in passing, however, that the use of Ahlstrom's (1943) rate-of-development schedule in calculating other sets of data may easily bring in systematic error if there be different relations between the temperature and the vertical distribution of pilchard eggs.

Unlike the estimates employing $\bar{c}'$, those using c'', $\bar{c}''$, c_a and c_b are entirely free from error connected with rate of development. This would be true not only for the present set of observations but for any set where the eggs are separable into daily age categories.

In general, therefore, the estimates of total eggs spawned in the survey area are negligibly affected by errors in the factors used for space and time weightings.

Errors Introduced by Interpolations. The estimates in this report assume continuity and linearity with space and time in the distribution of egg concentration. Both of these assumptions would tend toward validity as the spacing of stations in space and time becomes infinitesimal. The question of errors involved in these assumptions, therefore, becomes a question of whether or not the actually rather wide intervals in space and time completely destroy the validity of the assumptions or only introduce variability.

Two approaches are possible. In one, we may examine the effects of changes in spacing. In the other, we may examine evidence inherent in the patterns described by our sets of data.

Entering the first approach, it is, of course, impossible to experiment with closer spacing without collecting more material. Actually, a special set of material was collected in 1941, consisting of catches of pilchard eggs along a line of stations spaced 1½ to 3½ miles apart; hence, about 5–10 times more closely spaced than in the cruises here reported upon. Studies of the special set have not yet proceeded far enough to be included here in full. Gross inspection of the data, however, bears out the assumption of continuity. It suggests also that the assumption of linearity in the egg-concentration gradients between stations about 20 miles apart (as in the regular set) instead of the 1½ to 3½ miles apart (as in the special set) is materially but not disastrously faulty. The short-distance changes in the concentration gradient appear to be such as to be described better by curves than by straight lines. If the numbers of eggs per haul be plotted as ordinates on a base line scaled in miles, points connecting the tops of ordinates at about 20-mile intervals enclose a polygon whose area is not very dissimilar to that of the polygon constructed by connecting the tops of all ordinates. The material and its study are not sufficient to deduce any general relation between the spacing of stations and the egg numbers computed therefrom.

Although it is not possible, beyond this special set of data, to experiment with closer spacing of stations, we have experimented with wider spacing. For this purpose, the stations were divided into subsets by two procedures: (1) using alternate stations and (2) using alternate

TABLE III. Comparison of Estimates of the Number of Pilchard Eggs, in Billions, as Determined from Alternate Stations, from Alternate Lines of Stations, and from the Total Stations—All Estimates Making Use of the Formulation $C = D\Sigma (w'' \bar{c}_i'')$.

Sta. No.	*Total Sta.*	*Alt. Sta. I*	*Alt. Sta. II*	*Sta. Lines 1–3–5*	*Sta. Lines 2–4 & Adj. Sta.*
1940: 10	30220	22835	33576	19894	41836
11	28614	33989	23574	26750	31565
12	24704	26386	23023	16037	34609
13	62007	63236	60840	23300	107542
14	34377	32513	37954	30933	40453
15	3240	2996	3470	3676	2659
Survey Total, C_t	183162	181955	182437	120590	258664
Season Total, C_T	205800	204444	204985	135494	290634
1941: 17	77368	86960	67552	46239	114320
19	36452	22955	50700	23355	51862
21	13385	13269	13481	19734	5887
23	4425	4830	3999	3638	5352
25	256	418	108	298	197
27	22	18	46	15	49
Survey Total, C_t	131908	128450	135886	93279	177667
Season Total, C_T	158925	154759	163718	112384	214057

lines of stations, *i. e.*, lines 1, 3, and 5 in one group, lines 2 and 4 in the other. The estimate for each group was based on the average of the several age categories in each sample, $\bar{c}''$, and the stations were given equal weight in the integration over the area, w''. The results are given in Table III.

The two estimates for each season, based on groups of alternate stations, are nearly identical for the 1940 season and are within 5% of each other for the 1941 season. On the other hand, the estimates based on alternate lines of stations are quite dissimilar; the sum based on lines 1, 3, and 5 is only about half as large as that based on lines 2 and 4 in both 1940 and 1941. The high sum for the latter resulted from both lines passing through relatively high concentrations, whereas the low sum for the former resulted from only one of the three lines passing through the area of high concentration. Obviously, the spacing of lines much farther apart than the approximately 30-mile intervals of 1940 or 1941 would make a system of estimates based on interpolations relatively inaccurate.

Such empirical demonstrations pertain only to a particular set of data treated in a particular way and have little predictive value for results from other sets or from other possible treatments.

The other approach, that of examining evidence inherent in our set of data, may be of more general significance. First, as to continuity, inspections of charts for individual cruises, in which the distribution of eggs is expressed by drawing lines that are isometric with respect to egg concentration, show the irregularity of contours to be only slightly greater than the ones given in Figs. 3 and 4, which are for the means of the six cruises in 1940 and 1941, respectively. In all of them, the concentration changes in a more or less regular and gradual manner. There is always a zone (sometimes two) of high concentration, usually including several stations, with the concentration falling off in all, or nearly all, directions away from this zone.

There is not only evidence of this continuity but also some indication of the shape that a frequency surface would have, were one to erect at the station locations a set of ordinates proportional in height to the egg concentrations. Such a frequency surface would have relatively gentle contours in the zone of high concentration, with the contours falling off more sharply away from this zone and then becoming more gentle again when approaching the quite low concentrations at some distance from the zone of maximum and near much of the periphery of the survey area in our charts. Such a surface resembles a normal frequency surface in being convex upward near the mode, inflecting at some distance from the mode to become concave upward at a still greater distance. In fact, one may test the distribution of hauls, with respect to egg-counts, against the distribution expected from a normal frequency surface according to a method suggested by Sette (1943: 195–198). For the 1940 survey, such a test gives a probability of 0.57 that the observed hauls could have been drawn from a normal frequency surface or from a set of several such surfaces. Hence, an integration of egg numbers for our surface on the assumption of linear gradients would be analogous to representing a univariate frequency distribution by a frequency polygon, and our method of treating the station concentration as uniform in the "surface-area" (*i. e.*, uniform from the station to a point midway to adjacent stations) is analogous to a univariate frequency histogram. Since the areas enclosed by the frequency polygon and by the frequency histogram are identical when both are drawn through identical points, the frequency histogram will differ from the frequency curve by the same amount as does the frequency polygon. The area of the latter, when class intervals are relatively small, will not differ much from that of the frequency curve itself. However, for large class intervals, where, by chance, no class center is near the mode, a serious underestimate would result.

Translated to an appraisal of the situation in our set of data, this means that our estimates are in error, through the assumption of linearity, if the stations were so far apart that none chanced to lie at or very near the point of maximum concentration. This would produce a systematic error tending toward lowering the estimate, first, because the maximum values contribute most to the sum, and second, because the normal frequency surface, which our set of data apparently simulates, has a much smaller area of above-average than below-average surface. Since it is unlikely that every cruise would have a station at or very near the point of maximum concentration, our estimates probably lie below the true values in both years. Furthermore, with stations spaced approximately the same distance apart in both seasons, the underestimation probably would be greater in the season in which the eggs were the less widely distributed. From Figs. 3 and 4, it would seem that the 1941 total would be the one that is more severely underestimated. We know of no presently available means of appraising the degree of underestimation and suggest that a study of the spacing of stations in relation to the distribution of eggs be given a high priority in future work in this field.

Comparison of charts of individual cruises, with isometric lines drawn with respect to egg concentration, with similar charts of temperature, salinity, oxygen content, phosphate content and dynamic surface, suggest that the zones of egg concentrations are located in a relatively stable relation to combinations of some aspects of these water qualities. It is suggested that efforts to relate the physical and chemical system to the distribution of eggs would not only be illuminating for the factors influencing spawning, but also would provide important material for designing cruise plans to improve quantitative estimates of spawning.

This discussion thus far has related exclusively to interpolations in space. Practically the same considerations and the same uncertainties exist with respect to time. Spawning may be expected to rise from low rates at the beginning of the season to a peak later in the season and then fall away to low values toward the end of the season. This seems true of the two seasons of the present set of data. With only six points to describe the course of rise and fall in each, there is a hazard of failure to chance upon near-peak values, similar to the hazard of failing to occupy stations in the area of near-maximal concentration, with the result that effects on the resulting estimates would be similar.

General Appraisal of Reliability of Estimates. It has been seen that sampling variability of hauls introduces an uncertainty of estimate of about 5% at the one-standard-deviation limit; measurements of time and space have introduced only negligible error; interpolations, due to the wide intervals between stations in space and cruises in time, have probably introduced a systematic error, perhaps greater in one season than the other, but in both tending toward lowering the estimate.

In addition to this, there is another error systematically operating in the same direction, which so far has not been considered. That is, the effect of mortality during the incubation period. Our estimates were based on eggs in several age-categories. All groups, except eggs just spawned, have been subjected to some mortality, and therefore the estimates tend not to be of the number of eggs spawned but of the number of eggs part way through embryonic development.

The independent estimates of c_a and c_b throw some light on the probable amount of mortality and hence of the effect of mortality on our estimates. Since eggs of age-category "B" are one day older, on the average, than those of age-category "A," the former will have been subject to one day's more mortality than the latter. Therefore, one would expect the estimate of c_b to be lower than that of c_a. On the contrary, our estimate of c_b is larger than that for c_a in 1941, and the two are about equal in 1940. We can find no explanation for this, other than that the error of estimate is sufficiently great to obscure the effect of mortality. If so, the error due to the latter must be less than that due to the former and is probably negligible. Whether this probable negligibility extends to age-categories "C" and "D" is difficult to determine, because there are fewer data (due to incompleteness of daily age-categories), and random variability of the sum has larger play.

Finally, it should be noted that neither season's estimate embraced the whole spawning season in the survey area, and the survey itself included only a portion of the spawning range. Estimates of the percentage loss of eggs from our totals, due to incomplete coverage of the season within the survey area, have been given, but their reliability is undetermined. How far the estimate of totals for the survey area falls short of the total of spawning throughout the entire spawning range cannot even be estimated. Indeed, it is not determinable from our data whether the survey area included comparable parts of the total spawning activity in the two years. Hence, no conclusions may be reached as to the absolute or even relative amount of spawning in the two years, 1940 and 1941.

CONCLUSIONS

The study has brought to light some aspects of the problem of estimating spawning productivity, from quantitative plankton tows, that may be of significance to further advances in this field.

1. The estimation of total eggs in an area is possible by several methods, differing in detail, but when based on the same fundamental assumptions, the more easily computed ones give results similar to the more laborious ones for the particular station patterns employed in 1940 and 1941.

2. The station patterns were more nearly adequate in the spacing between stations on the station line than in the spacing between station lines, the latter not being susceptible to substantial increase without materially altering the estimates.

3. Although the sampling variability of our plankton hauls is large for individual hauls, it becomes moderate for the sum of such hauls.

4. Since the error connected with station spacing may be substantial and is probably systematic, the opportunity for advances in the understanding of factors influencing the distribution of the eggs is probably greater than that afforded by further refinements intended to decrease haul variability.

REFERENCES

Ahlstrom, E. H.

1943. Studies on the Pacific pilchard or sardine (*Sardinops caerulea*). 4.—Influence of temperature on the rate of development of pilchard eggs in nature. Spec. Sci. Rep. U. S. Fish Wildl. Serv., 23. 26 pp. (Mimeo.).

1948. A record of pilchard eggs and larvae collected during surveys made in 1939 to 1941. Spec. Sci. Rep. U. S. Fish Wildl. Serv., 54. 76 pp. (Mimeo.).

BUCHANAN-WOLLASTON, H. J.
1937. Two Technical Notes. (2) Estimating the total or the mean number of organisms under a given area of sea or a given area of sea bottom. J. Cons. int. Explor. Mer, *12* (3): 334.

CLARK, F. N.
1939. Measures of abundance of the sardine, *Sardinops caerulea*, in California waters. Fish Bull., Sacramento, 53. 45 pp.

HODGMAN, C. D.
1944. Hand book of chemistry and physics. 28th ed. Chemical Rubber Publishing Co., Cleveland. 2571 pp.

SCOFIELD, E. C.
1934. Early life history of the California sardine (*Sardina caerulea*), with special reference to distribution of eggs and larvae. Fish Bull., Sacramento, 41. 48 pp.

SETTE, O. E.
1943. Biology of the Atlantic mackerel (*Scomber scombrus*) of North America. Part 1. Early life history. Fish Bull., U. S. Fish Wildl. Serv., 38: 149–237.

SILLIMAN, R. P.
1946. A study of variability in plankton townet catches of Pacific pilchard (*Sardinops caerulea*) eggs. J. Mar. Res., *6* (1): 74–83.

SVERDRUP, H. U. AND STAFF
1944. Oceanographic observations on the "E. W. Scripps" cruises of 1940. Rec. Obser. Scripps Inst. Oceanogr., *1* (3): 161–248.
1947. Oceanographic observations on the "E. W. Scripps" cruises of 1941. Rec. Obser. Scripps Inst. Oceanogr., *1* (4): 249–407.

WINSOR, C. P. AND G. L. CLARKE
1940. A statistical study of variation in the catch of plankton nets. J. Mar. Res., *3* (1): 1–34.

WINSOR, C. P. AND L. A. WALFORD
1936. Sampling variations in the use of plankton nets. J. Cons. int. Explor. Mer, *11* (2): 190–204.

NEW GENERA AND SPECIES OF BATOID FISHES[1]

By
HENRY B. BIGELOW and WILLIAM C. SCHROEDER

Museum of Comparative Zoölogy
Harvard University
Cambridge, Massachusetts
and
Woods Hole Oceanographic Institution[2]
Woods Hole, Massachusetts

INTRODUCTION

The trawling campaigns carried out by the research vessel ATLANTIS of the Woods Hole Oceanographic Institution along the north and south coasts of Cuba during the winter of 1938 and spring of 1939, under the joint auspices of the University of Havana and of Harvard University, brought to light along the 200–500 fathom zone an abundant population of small skates (Family Rajidae), the existence of which had not previously been suspected there. Their discovery emphasizes the desirability of a more thorough exploration of the benthonic fish-fauna, especially for elasmobranchs of this general depth zone in the West Indian region as a whole, as well as at greater depths along the North American slope. The necessity of assigning these Cuban skates—representing several new species—to their correct taxonomic positions calls for some discussion of generic definitions within the large troublesome family to which they belong.

Some of them, with one other skate that has long been known, form a homogeneous group that is characterized by the shortness of the rostral projection from the cranium; a new genus, *Breviraja*, is proposed for these.

Two others of the new Cuban skates are of more general interest, for they illustrate an extreme modification of the pelvic fins. These fins, among elasmobranch fishes as a whole, diverge but little from what may be regarded as their primitive state, for they retain the broad base, have but little power of independent motion, and their outer margins are only moderately deeply concave, if at all so. However, each of the pelvics of these two Cuban skates is subdivided into

---

[1] Drawings by E. N. Fischer.

[2] Contribution No. 447 from the Woods Hole Oceanographic Institution.

two entirely separate lobes; the anterior lobe is slender, limb-like and arises independently from the lower surface of the disc, while the posterior one is much broader and fin-like, its anterior edge merging with the inner posterior margin of the corresponding pectoral fin. This modification of the pelvics is shared by two species of Rajidae from the Natal Coast of South Africa,[3] by the little-known genus *Anacanthobatis* Von Bonde and Swart (1924: 18, pl. 23) from Natal, and by one genus (*Typhlonarke*) of electric rays; otherwise it appears to be unique among batoids so far as known. Moreover, in these two Cuban species the modification of the pelvics is accompanied by a short extension rearward of each lateral bar of the pelvic arch itself beyond the point of articulations of the most posterior of the three or four radial cartilages borne by it; furthermore, the first two or three radials that are ordinarily borne along the anterior part of the basipterygial cartilage have been lost. The resultant gap (Fig. 1) corresponds to the external gap between the two subdivisions of the pelvic fin. Thus, the anterior pelvic division includes only the three or four radials that arise directly from the pelvis itself, whereas in all other skates, in which the condition in this regard is known, the anterior division includes these radials as well as the first two or three radials that are borne along the anterior part of the basipterygial cartilage. This specialization of the pelvic skeleton and fin is so remarkable that we have no hesitation in invoking it as the basis for a new genus, for which we propose the name *Cruriraja*, roughly translatable as "skates with legs."

A second major departure from the usual elasmobranch morphology is found in the conformation of the nostril of the little electric ray from the Pacific Coast of Central America, described as *Discopyge ommata* Jordan and Gilbert (1890). In all other electric rays, so far as known, the nostril is a simple aperture, more or less narrowed but not interrupted in its midsector. We have checked this feature in *Torpedo*, in *Narcine*, in *Hypnos*, in *Benthobatis* and in *Narke*. Such is the usual state of the nostril among the batoids as a whole, and no essential departure from this condition seems to have been reported previously for any elasmobranch. However, examination of a female and a male of *D. ommata*, loaned us by John Tee-Van of the New York Zoological Society, shows that each of the nostrils is divided into two separate apertures by a cross bridge about midway of its length. The outer of the two apertures constitutes the exposed and visible part of the nostril, whereas the inner aperture is entirely roofed over by the joint curtain-

---

[3] *Raja parcomaculata* and *R. durbanensis* Von Bonde and Swart (1924: 9, 11, pl. 21, fig. 2, pl. 22, fig. 1).

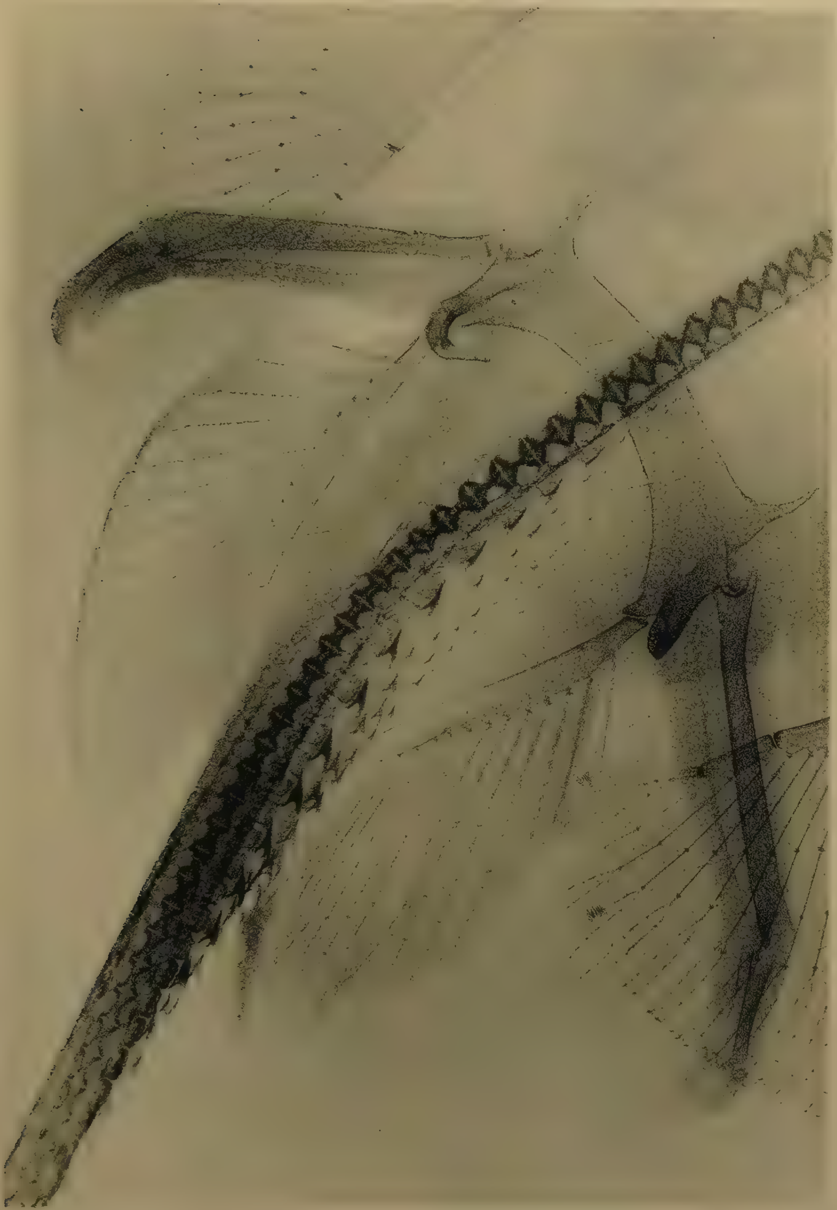


Figure 1. *Cruriraja poeyi,* sp. nov., female, type, 328 mm. long, Harv. Mus. Comp. Zool. No. 36324. X-ray photograph unretouched, showing the gap between radials borne on the pelvis and those on the anterior part of the basipterygial cartilage.

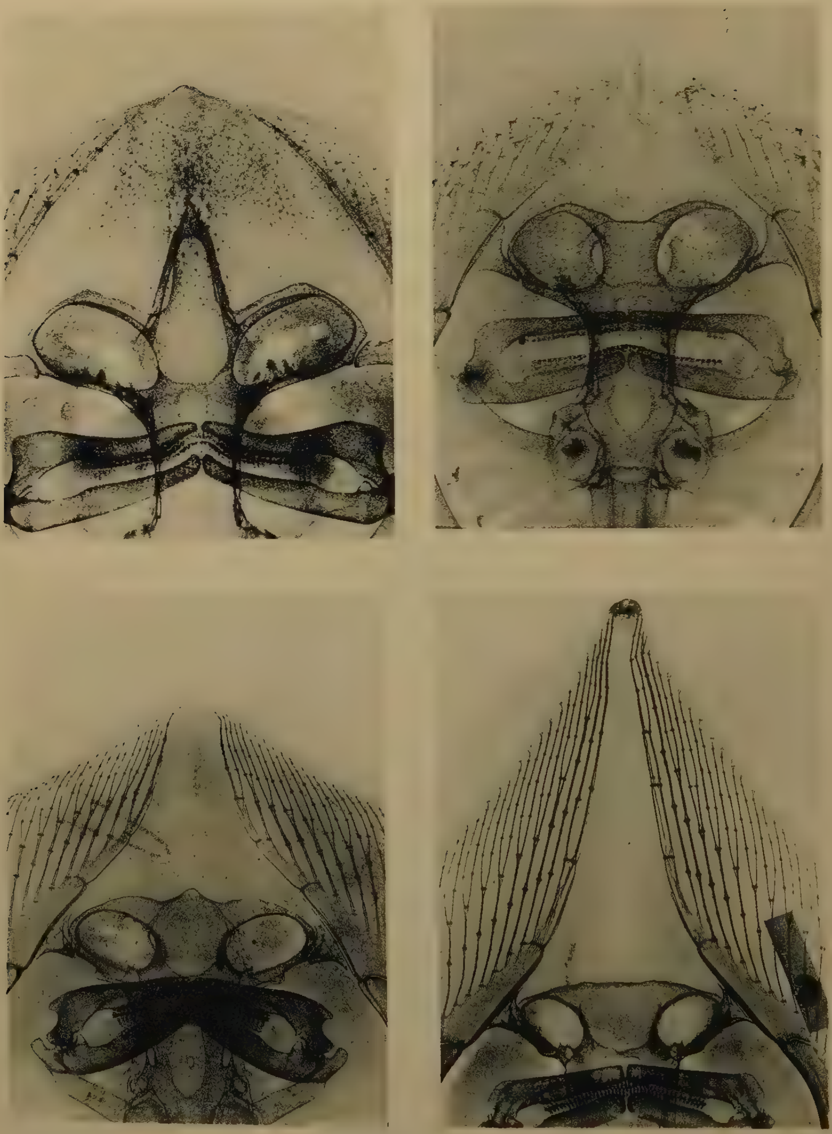


Figure 2. X-ray photographs, slightly retouched, showing front of cranium and rostral cartilage, where present, together with forward extension of pectoral rays. Upper left, *Breviraja colesi,* sp. nov. (Harv. Mus. Comp. Zool. No. 36374); upper right, *Psammobatis scobina* (Philippi, 1857), female, about 190 mm. long, from Uruguay (U. S. Nat. Mus. No. 86721); lower left, *Sympterygia microps* (Günther, 1880), female, about 475 mm. long, off Buenos Aires, Argentina (U. S. Nat. Mus. No. 55579); lower right, *Sympterygia bonaparti,* Müller and Henle (1841), female, from Argentina (U. S. Nat. Mus. No. 53437).

like expansion of the margin of the nostril and is wholly concealed unless the latter be lifted. This character, with the others listed below, sets *D. ommata* so sharply apart from all its relatives that it is necessary to establish for it a third new genus, *Diplobatis*.

FAMILY RAJIDAE

SKATES

Two genera of skates are so obviously set apart—the new *Cruriraja* by its pelvic fins and *Dactylobatus* Bean and Weed (1909) by the spatulate extension of the outer margin of each of its pectoral fins—that there is no danger of confusing them with any other known members of their family. The great majority of the remaining rajids—and they are many—fall in two categories as regards the degree of development of the rostral cartilage:

A. Those in which the rostral projection from the front of the cranium extends beyond the tips of the anterior radials of the pectoral fins, very nearly to the tip of the snout;

B. Those in which the rostral cartilage is either nonexistant or only very faintly indicated, and in which the snout region is supported by the forward extension of the pectoral rays on either side, aided more or less by ligamentous bands or sheets reaching forward from the anterior face of the cranium.

The species in category A (with rostral cartilage extending beyond pectoral rays) are so many, and those from different seas so closely resemble one another in many cases, that it would be a boon to students of fishes of the sea bottom if they could be grouped generically by any criteria that would be easy to see or to feel and at the same time give reasonably sharp alternative characters. However, all of the characters that have been proposed for this purpose, such as length and shape of snout, presence or absence of caudal fin-membrane, relative positions of dorsal fins on the tail, spination, and structures of claspers in adult males, have proved to intergrade so completely that there seems no escape from referring all the known members of this group to the old genus *Raja* of Linnaeus (1758).

The members of category B (*i. e.*, either those in which the rostral projection from the cranium is lacking altogether, [Fig. 2, upper and lower right] or in which it is represented by a very slight elevation only [Fig. 2, lower left]) are perhaps separable with reasonable sharpness into two genera, *Sympterygia* Müller and Henle (1841) and *Psammobatis* Günther (1870), according to the degree of concavity of the outer margins of the pelvics, as indicated in the following key. But critical intercomparison, especially of longer known species with those recently described from Peru (Hildebrand, 1946: 82), may prove that complete gradation exists between these extremes, in which case the older name *Sympterygia* will have precedence.

The Cuban collection now reveals the existence of a third category of typical skates, which includes several small species in which the rostral cartilage, while definitely developed, falls short of the tips of the pectoral rays (Fig. 2, upper left) and hence falls shorter still of the tip of the snout. Examination of the Museum of Comparative Zoology collections shows that *Raja plutonia* Garman (1881, 236; 1913: 335, pl. 18, fig. 1) from the upper part of the Continental Slope in the offing of South Carolina also falls in this category in respect to its rostral cartilage. In fact, Garman, in his original account of this species, remarked that the rostral cartilage did not extend to the end of the snout.

Furthermore, we have yet to see a skate, large or small, fresh or preserved, or from whatever ocean, that cannot readily be referred to one or the other of these three categories. The presence or absence of a rostral cartilage, and its length relative to the pectoral rays if present, is easily determined by touch and, in most cases, by viewing it against a strong light; such examination is possible even for large specimens, while x-ray photographs are even more revealing. Such being the case, we now propose for this category with short rostral cartilage the new genus *Breviraja*.

KEY TO GENERA OF RAJIDAE

1a. Middle rays of pectorals prolonged as spatulate extensions of disc. *Dactylobatus* Bean and Weed (1909).

1b. Middle rays of pectorals not prolonged.

  2a. Anterior lobes of pelvics in the form of slender, limb-like structures, entirely separate, externally, from the posterior lobes, and arising independently from the lower surface of disc. *Cruriraja*, gen. nov., p. 549.

  2b. Anterior lobes of pelvics fin-like (as are the posterior lobes) and continuously connected with the posterior lobes along outer margin of fin.

    3a. Anterior face of cranium with a definite rostral process, longer or shorter.

      4a. Tip of rostral process extending forward beyond level of tips of anterior rays of pectorals very nearly to extreme tip of snout. *Raja* Linnaeus (1758).

      4b. Tip of rostral process falling short of level of tips of anterior rays of pectorals, and falling shorter still of extreme tip of snout. *Breviraja*, gen. nov., p. 558.

    3b. Anterior face of cranium without a definite rostral process, or at most with only a very slight elevation, or even concave in some cases.

5a. Outer margins of pelvics so deeply concave that the fins are divided. *Psammobatis* Günther (1870).[4]

5b. Outer margins of pelvics only very weakly concave, the fins not divided. *Sympterygia* Müller and Henle, 1841.

Cruriraja,[5] GEN. NOV.

Generic characters. Rajidae with outer margin of pelvic so deeply notched that the anterior division is entirely cut off from the remainder of the fin, thus forming a separate limb-like structure that arises independently from the lower surface of the disc some little distance inward from the edge of the disc; inner posterior margin of pectoral continuous with anterior margin of posterior division of pelvic. Anterior division of pelvic subcylindrical, slender, tapering toward the tip, stiff proximally but softer distally, of three articulated segments, with the stoutness of its musculature toward the base suggesting considerable powers of movement; its support is given by the first stout radial and the succeeding two or three slender radials, all of which are articulated to the pelvic arch. Posterior lobe of pelvic fin-like, supported by about 18 radials only, the gap resulting from the loss of the anterior two or three radials that are ordinarily borne along the anterior part of the basipterygial cartilage corresponds in position to the interspace seen, externally, between the anterior and posterior subdivisions of the pelvic fin. Characters otherwise as in *Raja*.

The function of the limb-like specialization of the anterior part of the pelvic fin is not known. It may be tactile, as are the feeler-like pelvics of the hakes of the gadoid genus *Urophycis*, or it may assist in the progress of its owner over the bottom.

Type species. *Cruriraja atlantis* Bigelow and Schroeder (p. 550), the type specimen, Harv. Mus. Comp. Zool., No. 36320; ATLANTIS Station 3468, Lat. 23° 11′ N, Long. 81° 23′ W, off the north central coast of Cuba; 375 fathoms.

KEY TO KNOWN SPECIES OF *Cruriraja*

1a. Interspace between 1st and 2nd dorsal fins at least 1.5 times as long as base of 1st dorsal. *atlantis*, sp. nov., p. 550.

1b. Interspace between 1st and 2nd dorsal fins not more than about half as long as base of 1st dorsal.

---

[4] *Malacorhina* Garman (1877) is clearly a synonym of *Psammobatis*, as pointed out by Regan (1914: 22) and by Norman (1937: 33). This appears to apply equally to *Irolita* Whitley, 1931.

[5] From the latin *crus*, a leg or limb, and *Raja*, the generic name of the typical skates.

2a. Large thorns on head and from posterior part of disc out along tail, but no large thorns along midbelt of back between levels of spiracles and outer corners of pectorals. *poeyi*, sp. nov., p. 555.

2b. Large thorns in an unbroken series along midline of back from close behind level of spiracles rearward along disc and tail.

  3a. A group of thorns on tip of snout, others along anterior half of rostral ridge, and two thorns in the interspace between the two dorsal fins. *parcomaculata* Bonde and Swart (1924). Off Natal Coast, S. Africa; 298 fathoms.

  3b. No thorns on tip of snout, along rostral ridge, or in the interspace between the two dorsals. *durbanensis* Bonde and Swart (1924). Off Natal Coast, S. Africa; 420 fathoms.

Cruriraja atlantis, sp. nov.

Figures 3, 4

Distinctive Characters. The complete subdivision of its pelvic fins separates *C. atlantis* from other known members of the family Rajidae except for its genus mates *C. poeyi*, *C. parcomaculata* and *C. durbanensis;* and the great width of the interspace between its first and second dorsal fins marks it off from all three of these species. The fact that the upper surface of its disc is nearly uniformly covered with small prickles further separates it from *poeyi;* the rounded outer corners of its disc distinguish it from *parcomaculata;* and the several rows of enlarged thorns on its tail separate it from *durbanensis*, which has a single row from nuchal region to first dorsal when young and a single row falling short of first dorsal in older examples.

Description. Proportional dimensions in per cent of total length. Female, type, 332 mm. long, with disc 175 mm. wide. Harv. Mus. Comp. Zool. No. 36320.

Disc: length 40.7; breadth 52.7.

Snout length in front of: orbit 9.5; mouth 12.3.

Orbits: longest diameter 5; breadth between 2.6.

Spiracles: length 2.6; breadth between 6.3.

Mouth: breadth 5.6.

Nostrils: breadth between 5.1.

Gill openings: distance between 1st 11.9; 5th 6.6; lengths 1st 1.1; 3rd 1.3; 5th 1.0.

Distance: snout to cloaca (middle) 38.0; cloaca (middle) to tip of tail 62.0.

First dorsal fin: vertical height 2.4; length of base 3.6.

Second dorsal fin: vertical height 1.8; length of base 3.8.

Interspace between: 1st and 2nd dorsals 9.0; 2nd dorsal and tip of tail 4.4.

Pelvics: length of anterior margin 9.0.

Disc 1.3 times as broad as long, its anterior angle (tip of snout to level of anterior margins of orbits) about 94°, the anterior margins very slightly sinuous, weakly convex opposite orbits; the outer posterior and inner margins gently convex, the outer corners narrowly rounded, the axis of greatest breadth 70% of distance rearward from tip of snout toward axils of pectorals. Rostral cartilage narrow, extending nearly to tip of snout. Pectoral rays terminating anteriorly halfway between orbits and tip of snout. Snout sharp-tipped, its length in front of orbits 1.9 times as long as orbit, its length in front of mouth 2.4 times as long as distance between nostrils. Orbit twice as long as interorbital space and as length of spiracle. Mouth very slightly curved, without a noticeable bow in center. Posterior margin of nasal curtain weakly and irregularly fringed, the outer posterior margin of nostril smooth. Distance between inner ends of 5th pair of gill openings 1.3 times as great as distance between nostrils.

Teeth in 42 series in upper jaw, probably the same in lower, arranged in quincunx, with a low and rather blunt median cusp.

Anterior lobe of pelvic (entirely separated from posterior lobe) fleshy, limb-like, slightly longer than posterior lobe, with two radials besides the first stout one, which is of three segments. Posterior lobe continuous with pectoral, with 18 radials, its outer margin slightly wavy and moderately convex, inner margin straight or slightly concave. Tips of pelvics angular, extending back about ¼ the distance from axils of pectorals toward 1st dorsal.

Tail slender, 1.1 times as long from center of cloaca to origin of 1st dorsal as from tip of snout to cloaca, 1.4 times as long to 2nd dorsal, and 1.6 times as long to tip. First and 2nd dorsals similar in shape and about equal in size; interspace between 1st and 2nd dorsals 2.5 times as long as base of 1st dorsal. Base of 2nd dorsal about as long as base of 1st dorsal; distance from rear end of base of 2nd dorsal to origin of caudal fin-membrane about ½ as long as base of 2nd dorsal; distance from 2nd dorsal to tip of tail slightly greater than length of base of 2nd dorsal. A membranous fold, low down on side of tail, originates anterior to 1st dorsal by a distance about equal to base of latter, very narrow anteriorly, gradually widening rearward to maximum breadth opposite origin of 2nd dorsal and to tip of tail.

Entire upper surface of disc thickly and evenly set with small prickles, except for extreme tip of snout and a narrow band along outer posterior margin of each pectoral. Somewhat larger prickles

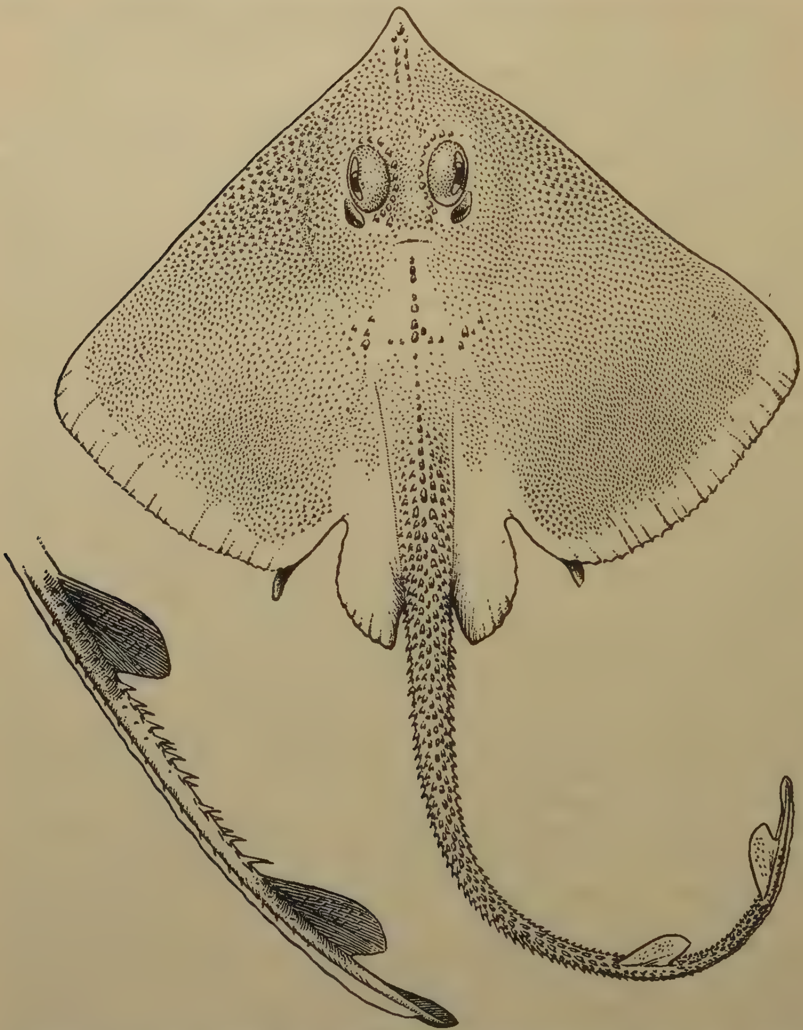


Figure 3. *Cruriraja atlantis*, sp. nov., female, type 332 mm. long (Harv. Mus. Comp. Zool. No. 36320), × 0.6. Dorsal fins and caudal fin membrane, × 1.2.

in the area opposite orbits, mostly with stellate bases so close together as to almost touch. Skin over eyes also prickly. Two or three rows of thorns on anterior part of rostral process. Inner margin of eye with an uninterrupted row of 11 thorns, besides 1 thorn posteriorly toward median line of back. Two or 3 low stout thorns on each scapular region. Midbelt of disc with a median row of 7 from nuchal region to pectoral girdle, followed by several very small thorns posterior to latter, increasing in size rearward. Also 1 or 2 additional rows of thorns on either side of middle row from a little in advance of axils

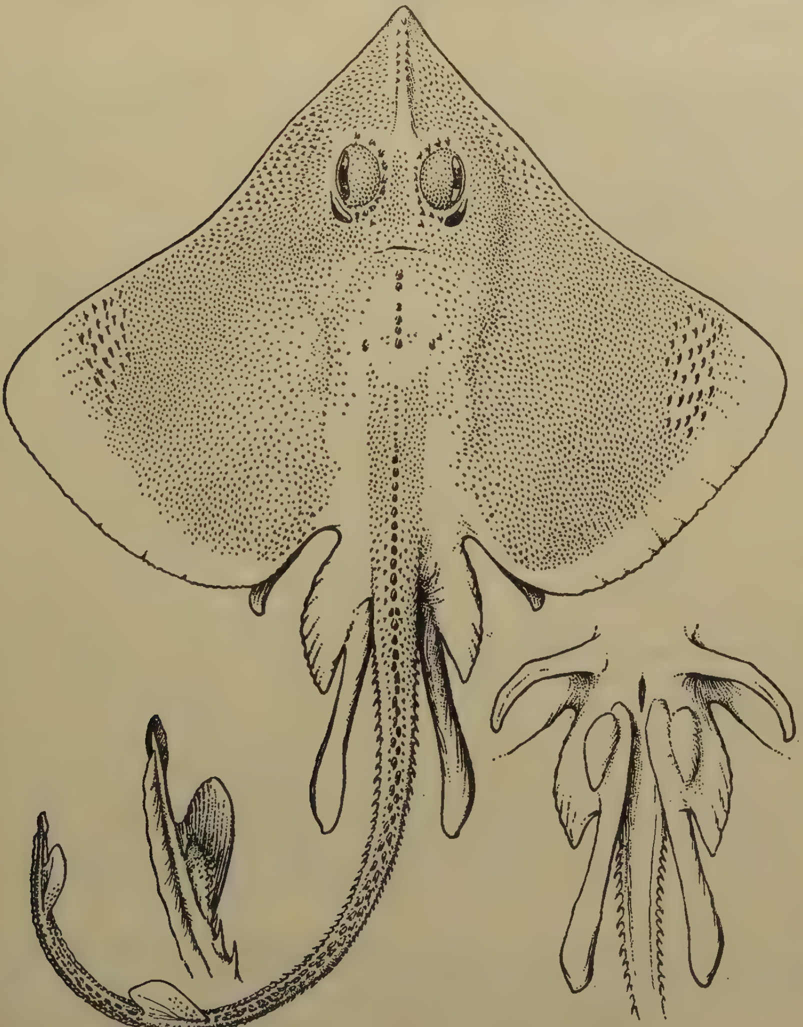


Figure 4. *Cruriraja atlantis*, sp. nov., male, 297 mm. long (Harv. Mus. Comp. Zool. No. 36321), × 0.7. Pelvic fins and claspers, × 0.7; second dorsal fin and caudal fin membrane, × 1.5.

of pectorals. Tail very thorny above and on sides, with 6 or 7 rows along anterior part, but with only 2 or 3 rows approaching the 1st dorsal; 1 or 2 rows between dorsals, succeeded by smaller prickles along sides of tail to a point opposite rear end of base of 2nd dorsal. Dorsals with a few prickles. Pelvics smooth. Lower surface of disc and tail smooth.

Color. Upper surface uniformly pale brown; disc without definite markings; a dusky band on the tail in the region of each dorsal fin; the dorsals black (a very conspicuous feature), as is the caudal. Lower surface whitish or yellowish, the abdomen and posterior parts of pectorals irregularly blotched with dark brown, the posterior lobes of the pelvics still darker brown.

Additional Specimens. A mature male, 297 mm. in total length and 161 mm. wide, with a small patch of alar thorns on each pectoral, differs from the female as follows: The anterior margins of the disc are concave opposite spiracles; the mouth is somewhat more curved; the outer posterior nasal margin has a few very short fringes; the posterior lobes of the pelvics are narrower and their tips more pointed; the tail has fewer thorns, for it lacks the rows immediately beside the median one. The lower surface of the disc is plain yellowish on the anterior half, pale dusky brownish on the posterior half, with an irregular demarkation separating the two areas; the anterior lobes of the lower surfaces of the pelvics are darker than in the female.

An immature male, 160 mm. long with disc 85 mm. wide, closely resembles the larger specimens except that the anterior margins of the disc are not noticeably concave opposite the spiracles. However, on a very young male, only 125 mm. long with disc 62 mm. wide, the pectorals are bare of prickles, although the midzone of the back is already prickly from the scapular region to opposite the tips of the pelvics; only 4 or 5 thorns have yet developed around the inner margin of each orbit. On the disc the median row of thorns is represented by only 2 or 3, these widely spaced, between nuchal region and level of axils of pectorals; these are succeeded by 1–2 rows of long sharp thorns rearward to the first dorsal, and between the dorsals. Color below plain whitish.

Developmental Stages. Nothing is known except that the new-born are about 100 mm. in total length.

Size. The largest specimens taken were a female 332 mm. long and a mature male 297 mm. Our collection contains 12 specimens, nine males and three females, 100 to 332 mm. in total length, with discs 50 to 175 mm. wide.

Habits. This is a deep-water species, all the specimens thus far having been taken in depths of 280 to 425 fathoms.

Range. Known only from off the north central coast of Cuba.

Cruriraja poeyi, sp. nov.

Figures 1, 5

Distinctive Characters. The pelvic fins of *C. poeyi* make it separable at a glance from all other members of the family Rajidae yet known from the western Atlantic, except for *C. atlantis.* It differs from the latter in the facts that the upper surface of its disc is largely smooth excepting around the eyes and snout and along the mid-dorsal zone posterior to the scapular region, and furthermore, its two dorsal fins are close together. The smoothness of the disc also separates it from *C. parcomaculata* and *C. durbanensis* from South Africa, which are more or less covered with small spines in addition to the enlarged median and orbital thorns.

Description. Proportional dimensions in per cent of total length. Female, type, 328 mm. long, with disc 208 mm. wide. Harv. Mus. Comp. Zool No. 36324.

Disc: length 48.2; breadth 63.5.
Snout length in front of: orbit 12.2; mouth 14.0.
Orbits: longest diameter 5.2; breadth between 3.3.
Spiracles: length 2.7; breadth between 7.9.
Mouth: breadth 7.0.
Nostrils: breadth between 6.4.
Gill Openings: distance between 1st 13.7; 5th 7.6; lengths 1st 1.4; 3rd 1.5; 5th 1.0.
Distance: Snout to cloaca (middle) 42.7; cloaca (middle) to tip of tail 57.3.
First dorsal fin: vertical height 1.8; length of base 4.0.
Second dorsal fin; vertical height 1.8; length of base 4.6.
Interspace between: 1st and 2nd dorsals 1.2; 2nd dorsal and tip of tail 4.0.
Pelvics: length of anterior margin 12.5.

Disc 1.3 times as broad as long, its anterior angle (tip of snout to level of anterior margins of orbits) 98°, the anterior margins slightly sinuous, being weakly convex opposite anterior part of orbits, the outer corners narrowly rounded, the posterior margins weakly convex, the posterior corners broadly rounded, the inner margins nearly straight, and the axis of greatest breadth 67% of the distance back from the snout toward axils of pectorals. Rostral process very narrow, reaching tip of snout, the anterior rays of pectorals falling well short. Snout pointed, its length in front of orbits 2.3 times as long as orbit, its length in front of mouth 2.2 times as long as breadth between nostrils. Orbit 1.6 times as long as distance between orbits and 1.9 times as long as spiracle. Mouth nearly straight. Nasal curtain fringed with a few very short knob-like lobes, the outer posterior margin of nostril smooth. Distance between inner ends of fifth gill openings 1.2 times as long as between nostrils.

Teeth in $\frac{46}{44}$ series, close set, in quincunx, circular at base, with a low triangular cusp pointing inward excepting near corners of mouth

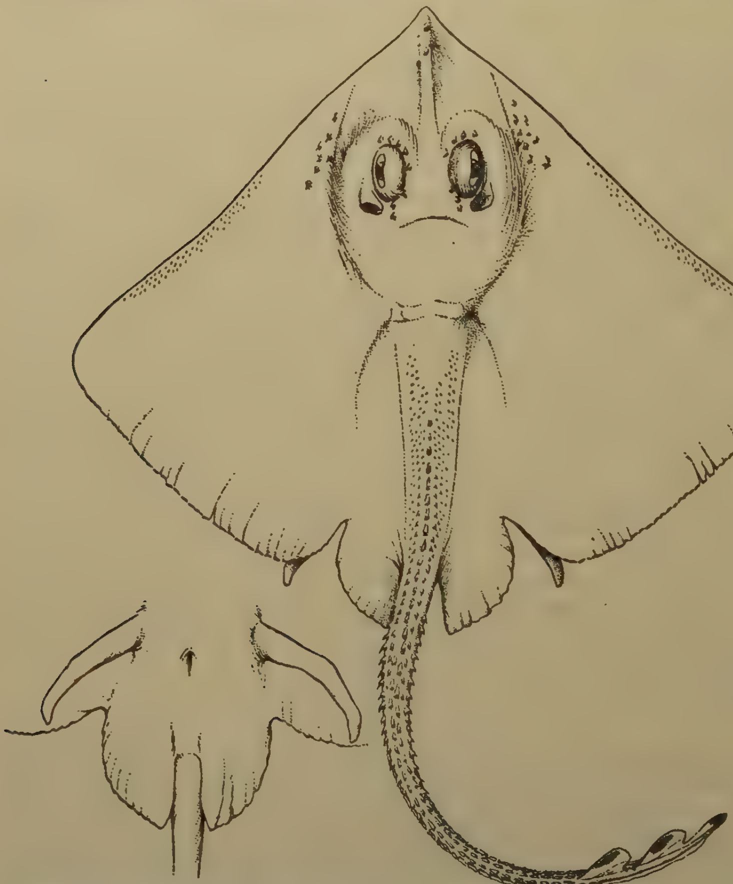


Figure 5. *Cruriraja poeyi*, sp. nov., female, type, 328 mm. long (Harv. Mus. Comp. Zool. No. 36324), × 0.5. Pelvic fins, × 0.5.

where cusps are usually lacking; the anterior rows worn nearly smooth.

Anterior lobe of pelvics about as long as posterior lobe. Outer margin of posterior lobe broadly rounded, its edge weakly scalloped, the inner margin nearly straight. Tips of pelvics moderately pointed, extending back about ¼ the distance from axils of pectorals toward 1st dorsal.

Tail narrow, its length from center of cloaca to 1st dorsal equal to length from snout to cloaca, 1.1 times to 2nd dorsal, 1.3 times to tip. Dorsals close together, the interspace about ⅓ as long as base of 1st dorsal, which is about equal to base of 2nd and to length of caudal fin. A narrow membranous fold, beginning low down on each side of tail, extends from about midway between tips of pelvics and origin of 1st dorsal to opposite the middle of caudal; widest opposite dorsals.

Upper surface of disc prickly over a narrow band along anterior margins from opposite the middle of orbits nearly to outer corner of disc, and along mid-dorsal ridge from a little posterior to pectoral girdle rearward. A small patch of thorns near end of rostral ridge, 7 or 8 between orbit and margin of disc; 8 on one side and 7 on the other along inner margin of orbit, the series interrupted about midway; also, a median row extending rearward from about midway between pectoral girdle and level of axils of pectorals, flanked by 2 or 3 rows of smaller thorns from a little anterior to axils of pectorals. Otherwise naked, including skin over eyes. Tail conspicuously rough; 6–7 rows of thorns anterior to tips of pelvics, although no one row is exactly in the midline, and 4 rows posteriorly, the anterior tail thorns more closely spaced than those rearward; 2 pairs of opposing thorns between dorsals. Dorsals prickly, but caudal and pelvics smooth. Lower surface smooth.

Color. Upper surface pale coffee brown except for a translucent area on either side of rostral cartilage; anterior parts of dorsals and tip of caudal membrane black. Lower surface of disc of about same shade as above, somewhat mottled with paler and darker, with pale areas around the mouth and along the gill openings. Tail paler than disc.

Additional Specimens. A male, 327 mm. in total length with disc 197 mm. wide, differs from the female chiefly as follows: Anterior angle of disc more acute (86°); snout sharper; posterior margin of nasal curtain conspicuously fringed on one side but almost smooth on the other; teeth in 50 series in upper jaw, with longer and sharper cusps; pelvics narrower, with tips more pointed; claspers well developed, extending beyond tips of pelvics a distance equal to width of mouth; a somewhat larger patch of malar thorns beside the orbits and a triangular patch of alar thorns on the outer parts of pectorals.

Several medium-sized specimens, 230 to 260 mm. long, with color better preserved, have roundish spots somewhat darker than the ground color; these are about half as large as the orbit and are scattered on the disc posterior to the spiracles, except along the outer posterior margins. A newborn specimen, 83 mm. in total length with disc 49 mm. wide, has a single median row of thorns from a little in advance of the axils of pectorals to the 1st dorsal. Seemingly these are lost with growth, and are replaced by one row on each side of the midline. There is also a single row low down on each side of the tail and a patch of prickles in the midzone from a little in advance of the axils of pectorals to a little posterior to the tips of the pelvics.

Developmental Stages. Nothing is known except that this skate, when newborn, is about 90 mm. in total length and closely resembles the adult.

Size. The largest specimen obtained was 328 mm. in total length. A male of about this size had well developed claspers, hence the total length reached by this species is probably not much greater than 400 mm. The collection contains 15 specimens, six males and nine females, 83 to 328 mm. long and 49 to 208 mm. wide; of these, 13 were trawled by the research vessel ATLANTIS and one by the U. S. Fisheries vessel ALBATROSS.

Habits. All specimens thus far known have been trawled in depths ranging from 210 to 475 fathoms, on hard sandy bottom.

Range. Known from off the north central and southwest central coasts of Cuba and off St. Augustine, Florida.

Breviraja, GEN. NOV.[6]

Rajidae with a rostral cartilage, but with the latter falling considerably short of the extremities of the anterior rays of the pectorals and hence short of the tip of the snout; the anterior pectoral rays of the two sides are either close together anteriorly or are farther separated. Characters otherwise as in *Raja*.

Type species, *Breviraja colesi* Bigelow and Schroeder, sp. nov.

Only one of the many species of the family previously described (*i. e.*, the *Raja plutonia* of Garman [1881]) is known with certainty to fall in this new genus. But the Cuban collections of the ATLANTIS include several new species that belong here, one of which is described below as type of the genus; the description and naming of the others will be postponed until a later occasion.

---

[6] From the latin "brevis," and *Raja*, the generic name of typical skates.

Breviraja colesi, sp. nov.

Figures 2 (upper left), 6

Distinctive Characters. Two characters, namely the lack of enlarged thorns in the mid-dorsal belt of the disc between the scapular region and the pelvic girdle, and a coloration which includes brownish roundish spots (some spots on the adult having pale centers), set off *B. colesi* from *B. plutonia* (Garman 1881), the only other skate yet described from the western Atlantic with pectoral rays extending forward beyond the rostral process.

Description. Proportional dimensions in per cent of total length. Female, type, 333 mm. long, and 160 mm. wide. Harv. Mus. Comp. Zool. No. 36374.

Disc: length 45.9; breadth 48.0.

Snout length in front of: orbit 9.8; mouth 11.2.

Orbits: longest diameter 4.5; breadth between 3.0.

Spiracles: length 2.8; breadth between 5.7.

Mouth: breadth 6.3.

Nostrils: breadth between 6.6.

Gill openings: distance between 1st 13.6; 5th 6.6; lengths 1st 1.8; 3rd 1.8; 5th 1.5.

Distance: snout to cloaca (middle) 41.4; cloaca (middle) to tip of tail 58.6.

First dorsal fin: vertical height 1.5; length of base 3.6.

Second dorsal fin: vertical height 1.5; length of base 3.6.

Interspace between: 1st and 2nd dorsals 0.0; 2nd dorsal and tip of tail 1.5.

Pelvics: length of anterior margin, 10.2.

Disc 1.05 times as broad as long, its anterior angle (snout to level of anterior end of rostral process) 135°, the anterior margins, outer corners, outer posterior margins, posterior corners and inner margins of disc all rounded, the outer corners most broadly so, the axis of greatest breadth 70% of the distance back from tip of snout toward axils of pectorals. Rostral process short, triangular, forming an angle of 40° (x-ray), extending rigidly for only ⅔ of the distance from center of mouth toward tip of snout. Pectoral rays extending forward well in advance of rigid rostral process. Snout blunt, with a small protuberance at tip, its length in front of orbits 2.16 times as long as orbit, its length in front of mouth 1.7 times as long as distance between nostrils. Orbit 1.5 times as long as distance between orbits and 1.6 times as long as spiracle. Mouth bowed centrally. Posterior margin

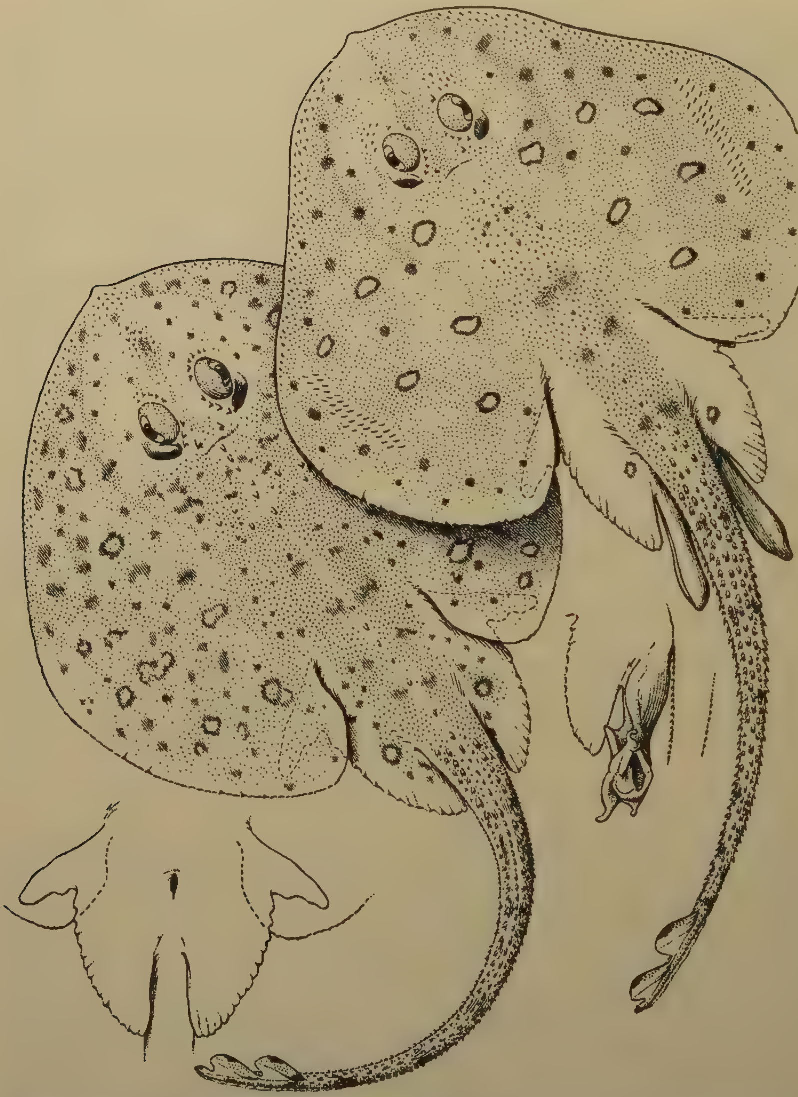


Figure 6. *Breviraja colesi*, sp. nov., female, type, 333 mm. long, and pelvic fins (Harv. Mus. Comp. Zool. No. 36374). Male, 320 mm. long (Harv. Mus. Comp. Zool. No. 36447); and clasper of a mature male; all × 0.5.

of nasal curtain prominently fringed, the outer posterior margin of nostril finely so. Distance between 5th gill openings equal to distance between nostrils.

Teeth in 50 series in upper jaw, probably about the same in lower, close set, in quincunx, with a very small cusp on an ovate base.

Pelvics deeply notched, the anterior lobe with two radials besides the first stout one, which has three segments, its posterior margin slightly wavy, its anterior margin about ⅔ as long as distance from origin of pelvic to rear tip of lobe; the posterior lobe with 24 radials, its outer margin weakly scalloped and moderately convex, its inner margin nearly straight. Tips of pelvics narrowly rounded, extending back about ⅜ of the distance from axils of pectorals toward 1st dorsal.

Tail slender, its length from center of cloaca to origin of 1st dorsal 1.2 times as long as from tip of snout to cloaca, its length to 2nd dorsal 1.3 times and to tip 1.4 times. Dorsals confluent, the base of 1st dorsal as long as base of 2nd. Length of caudal membrane about ⅖ as long as base of 2nd dorsal. A thin and narrow lateral fold, originating on each side of tail a little beyond tips of pelvics, widens somewhat posteriorly and continues to extreme tip.

Upper surface of disc, including skin over eyes, rough with small close-set prickles except at margins around posterior angles where they are sparse. A narrow band of very small thorns (slightly larger than the prickles) along anterior margin from level of spiracles forward; a patch of small thorns over end of rostral process about ⅔ of distance from orbits toward tip of snout; inner margin of orbit to spiracle with 10 or 11 thorns, the row interrupted opposite middle of orbit; a median pair of thorns on nuchal region a little posterior to spiracles; 4 or 5 small thorns on each scapular region. An interrupted row of small thorns along midline of back—6 from the nuchal region to pectoral girdle, 6 more widely spaced to the level of the axils of pelvics, followed by 44 more closely crowded and less regularly arrayed along the tail to the origin of the 1st dorsal. Each side of tail with 2 or 3 irregular rows of larger thorns that become increasingly crowded rearward, and with smaller thorns extending from the level of 1st dorsal to the tip. Dorsals prickly on anterior half or ⅔ of fins. Pelvics with a few small prickles on central part of posterior lobe. Disc and tail smooth below.

Color. Upper surface of disc pale brown mottled with numerous irregular spots and blotches of darker brown, its outer part also with a variable number of larger roundish spots with pale centers and brown margins, one being situated on the center of each posterior pelvic lobe. Tail with various pale brown markings, some appearing as faint crossbars, its lateral membranous folds whitish. Upper anterior part of each dorsal and anterior base of 1st dorsal with a bronze blotch. Lower surface of disc and of tail pale yellowish or whitish.

Additional Specimens. An immature female 150 mm. long has only 5 or 6 thorns around the inner margin of the orbit, 3 median thorns from the nuchal region to the pectoral girdle, and one only over each scapular region; the thorns on the tail are in more regular rows than in the adult female, and there are 7 or 8 obscure cross bars on the tail. A mature male 320 mm. long has the anterior margin of the disc slightly concave opposite and posterior to the spiracles, the outer angle of the pectorals are more broadly rounded and the tips of the pelvics are more narrowly rounded than on the female; also the tail is longer (1.6 times as long from center of cloaca to tip as from snout to cloaca). The outer margins of the disc opposite and in front of the orbits are armed with malar thorns, slightly larger than the prickles, and there are 2 to 3 rows of alar spines near the outer angles of the pectorals, the extreme length of the rows being a little greater than the distance between the outer ends of the spiracles. The male resembles the female in general coloration, but it has fewer blotches and spots and lacks the blotch at the base of the first dorsal.

Developmental Stages. Nothing is known except that a very young specimen, with disc only 43 mm. wide, bears a close resemblance to the adults.

Size. This species apparently does not reach a length of more than 400 mm., the largest female taken being 333 mm. long, while a mature male was only 320 mm. We have eight specimens of this species, four males and four females, 80 to 333 mm. in total length, with discs 43 to 160 mm. wide, all trawled by the research vessel ATLANTIS.

Habits. All specimens thus far known have been trawled in depths ranging from 200 to 285 fathoms.

Range. Known from Cochinas Bay on the south coast of Cuba, and off Matanzas and Santa Clara Provinces of the north coast.

FAMILY TORPEDINIDAE

ELECTRIC RAYS WITH TWO DORSAL FINS

Diplobatis, GEN. NOV.[7]

Figures 7, 8 (right), 9

Generic characters. Torpedinidae with each nostril subdivided about midway of its length into two separate apertures by a cross bridge of stiff tissue; outer aperture exposed but the inner entirely roofed over by the joint nasal curtain. *Diplobatis* falls with *Narcine* in most of its other characters; *i. e.*, narrow protractile mouth with labial cartilages at the corners, normally developed eyes with rather prominent orbits, relatively long tail with lateral folds, pelvics entirely separate one from the other and free from the sides of the tail at their extreme tips, and nasal curtain much wider than long. However, the relationship of its tooth bands to its thick, fleshy lips affords a second generic charac-

---

[7] διπλοος, double, referring to the structure of the nostril, and βατις, a skate or ray.

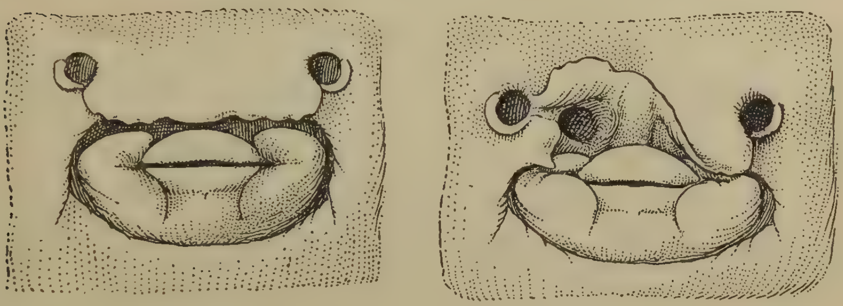


Figure 7. *Diplobatis ommata*, female, about 98 mm. wide; left, mouth and nostril with nasal curtain lifted only enough to show shape of upper lip; right, with right-hand side of nasal curtain rolled back to show the transverse division of the nasal aperture. × 2.75.

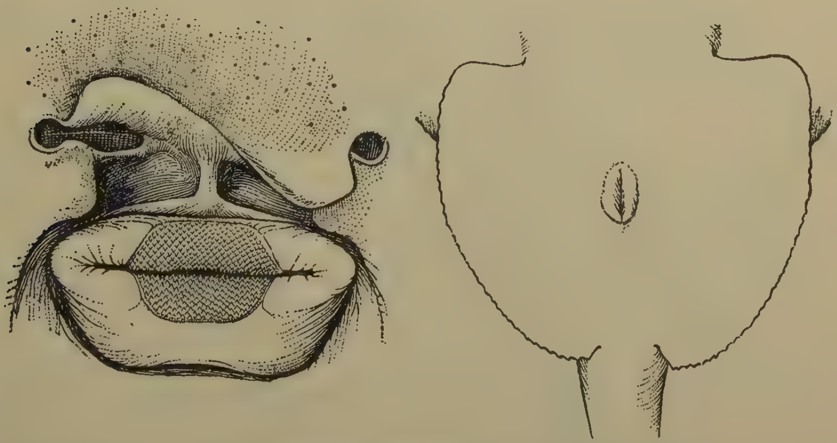


Figure 8. Left, *Narcine brasiliensis*, specimen 218 mm. wide, mouth and nostril with nasal curtain rolled back to show single nasal aperture, and extension of dental bands over the lips, × 1.33; right, *Diplobatis ommata*, ventral view of pelvic fins of female, about 98 mm. wide, × 0.7.

ter; its teeth are entirely concealed when the mouth is retracted and closed, while the median sector of each lip, corresponding in its sidewise extent to the sidewise extension of the tooth bands within the mouth, is marked off at either end by a deep transverse groove when

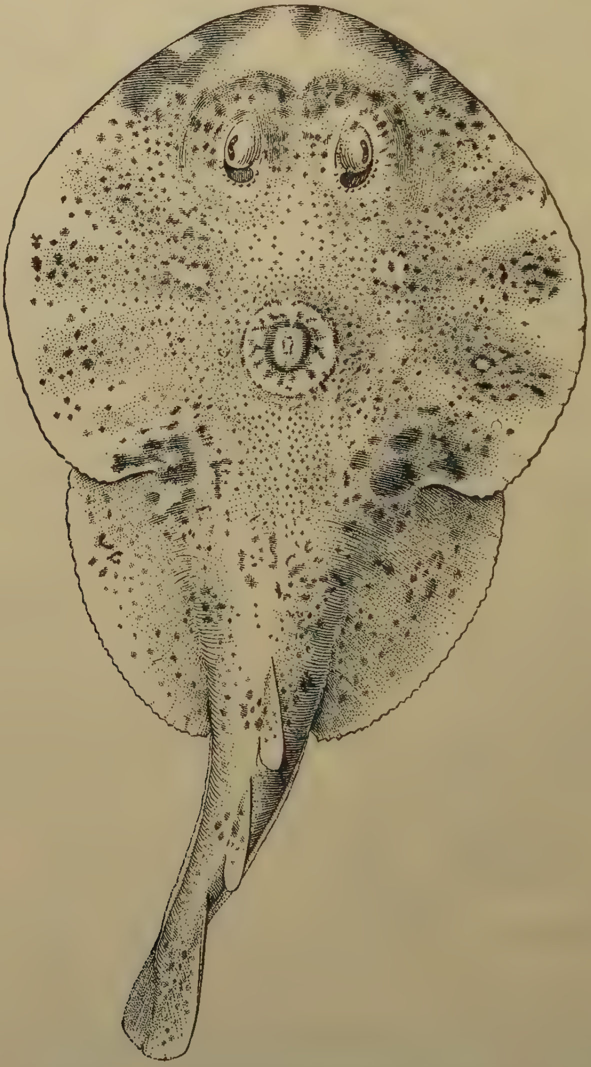


Figure 9. *Diplobatis ommata*, female, about 98 mm. wide, to show color pattern.

the mouth is retracted. In all other genera of the family in which the mouth is protractile,[8] the tooth bands extend so far forward and the median sectors of the lips are so thin (anteroposteriorly) that several of the anterior rows of teeth are fully exposed even when the mouth is tightly closed. And in them the lips are not crossed by definite transverse grooves corresponding with the lateral limits of the tooth bands, even when the mouth is fully retracted (Fig. 8). The type species is *Discopyge ommata* Jordan and Gilbert (1890), of which *Narcine ommata* Clark (1936) appears to be a synonym; Pacific Coast of America, Gulf of California to Panama.

The color pattern of this interesting little electric ray is so conspicuous and so characteristic that we include a drawing to show it (Fig. 9); see Beebe and Tee-Van (1941, pl. 1, Fig. 2.) for a photograph showing a very similar pattern for a male specimen.

REFERENCES

Bean, B. A. and A. C. Weed
1909. Description of a new skate (*Dactylobatus armatus*) from deep water off the southern Atlantic Coast of the United States. Proc. U. S. nat. Mus., *36:* 459–461, pl. 38.

Beebe, C. William and John Tee-Van
1941. Eastern Pacific Expeditions of the New York Zoological Society. XXVIII. Fishes from the tropical eastern Pacific. [From Cedros Island, Lower California, south to the Galápagos Islands and northern Peru]. Part 3. Rays, mantas and chimaeras. Zoologica, N. Y., *26:* 245–280.

Clark, H. W.
1936. The Templeton Crocker Expedition of the California Academy of Sciences, 1932. New and noteworthy fishes. Proc. Calif. Acad. Sci., (4) *21:* 383–396.

Garman, S. W.
1877. On the pelvis and external sexual organs of selachians, with especial references to the new genera *Potamotrygon* and *Disceus* (with descriptions). Proc. Boston Soc. nat. Hist., *19:* 197–215.
1881. Reports on the results of dredging, under the supervision of Alexander Agassiz, along the Atlantic Coast of the United States. Report on the selachians. Bull. Mus. comp. Zool. Harv., *8* (11): 231–237.
1913. The Plagiostomia (sharks, skates, and rays). Mem. Harv. Mus. comp. Zool., *36:* i–xiii + 1–515; Atlas, 75 pls.

Günther, Albert
1870. Catalogue of the fishes in the British Museum. Taylor and Francis, London, *8.* xxv–549 pp.

Henle, F. G. J.
1834. Ueber Narcine. G. Eichler, Berlin. 44 pp., 4 pls.

---

[8] Genera *Narcine, Benthobatus, Discopyge,* and *Heteronarce* (probably).

HILDEBRAND, S. F.
1946. A descriptive catalog of the shore fishes of Peru. Bull. U. S. nat. Mus., 189: 1–530.

JORDAN, D. S. AND C. H. GILBERT
1890. In D. S. Jordan and C. H. Bollman: Scientific results of explorations by the U. S. Fish Commission Steamer Albatross. IV. Descriptions of new species of fishes collected at the Galápagos Islands and along the coast of the United States of Colombia, 1887–88. Proc. U. S. nat. Mus., *12:* 149–183.

LINNÉ, CARL VON
1758. Systema naturae per regna tria naturae . . . editio decima, reformata. Laurentii Salvii Holmiae. *1.* 824 pp.

MÜLLER, JOHANNES AND JOKOB HENLE
1841. Systematische Beschreibung der Plagiostomen. Von Veit und Comp., Berlin. xxii–200 pp., 60 pls.

NORMAN, J. R.
1937. Coast fishes. Part II. The Patagonian region. 'Discovery' Rep., *16:* 1–150, pls. 1–5.

REGAN, C. T.
1914. Fishes. Nat. Hist. Rep. Brit. Antarc. ("Terra Nova") Exped. 1910–1913. Zool., *1:* 1–54, pls. 1–13.

VON BONDE, C. AND D. B. SWART
1924. The platosomia (skates and rays) collected by the S. S. "Pickle." Rep. Fish. Mar. Biol. Surv., S. Afr. (1922) 3, Spec. Rep. 5: 1–22.

WHITLEY, G. P.
1931. Studies in ichthyology. 4. Rec. Aust. Mus., *18* (3): 96–133, pls. 11–16.

A BIOCHEMICAL STUDY OF THE MARINE ANNELID WORM, *THORACOPHELIA MUCRONATA*

ITS FOOD, BIOCHROMES AND CAROTENOID METABOLISM [1, 2]

By

DENIS L. FOX, SHELDON C. CRANE AND
BAYARD H. McCONNAUGHEY[3]

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

Comparative biochemical studies of numerous marine invertebrates and fishes have indicated that the majority of such species so far examined selectively assimilate and store xanthophyllic or allied oxygenated carotenoids rather than carotenes, when consuming food containing both types of pigment. In some forms there is complete exclusion of carotenes, *e. g.*, in several fishes, in the sea mussel (*Mytilus californianus*), in at least three brittle stars, and in some color-variants of the anemone *Metridium senile.* Some asteroid echinoderms store carotenes, but in far lower concentrations than xanthophylls, while four species of echinoids appear to assimilate relatively greater quantities of carotenes (Fox, Updegraff and Novelli, 1944).

The marine polychaete worm *Thoracophelia mucronata* represents a metabolic departure from the above groups, since, while consuming a basic diet of finely particulate marine detritus similar to that ingested by mussels, brittle stars and sand dollars, it apparently stores only the carotene type (Sumner and Fox, 1933; Fox, 1936) and seems to destroy much of the xanthophyllic material. This is in direct contrast to *Mytilus* (Scheer, 1940) and to the ophiuroids *Ophiopteris papillosa, Ophiothrix spiculata* and *O. rudis*, which store xanthophylls but not carotenes (Fox and Scheer, 1941). *Thoracophelia* also exhibits more complete xanthophyll exclusion than does the sand dollar, *Dendraster*

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 378.

[2] This research was aided in part by a grant from the Penrose Fund of the American Philosophical Society and more recently by a grant from The Rockefeller Foundation. To the officers of each Institution grateful acknowledgment is made.

[3] The authors take pleasure in acknowledging the technical assistance of Mr. Arthur S. Lockley, and helpful information from Professor Wesley R. Coe.

excentricus, for example, whose body carotenoids are some 24% xanthophylls. Other echinoids likewise exhibit a comparable preponderance of carotenes (Fox and Scheer, 1941).

Reference should be made also to the sponges. For while these must ingest much organic detritus, certain of them further resemble *Thoracophelia* in assimilating carotenes with relative exclusion of xanthophylls. *Suberites domuncula* and *Ficulina ficus* seem to yield no acidogenic carotenoids, questionable amounts of xanthophylls and large quantities of persistently epiphasic carotenoids with spectra resembling those of torulene, lycopene (*in Suberites*), α-, β-, and γ-carotenes (Lederer, 1938). The red sponge *Hymeniacidon sanguineum* was found to contain α-, β-, and γ-carotenes, as well as the ketonic carotenoid echinenone, more commonly encountered in sea urchins (Drum and O'Connor, 1940; Drumm, O'Connor and Renouf, 1945).

Because of its vast numbers, *Thoracophelia mucronata* is an undoubted factor in the physical and chemical turnover of organic matter on local sandy shores, in a manner somewhat analogous to the function of earthworms in terrestrial soils. Furthermore, this species has long served, and is still employed, as an important and economical source of food for experimental marine animals, including some collections maintained upon diets designed to provide carotene. Finally, the species constitutes an interesting case of an unusual biochemical ability—that of absorbing only carotene, with seemingly complete exclusion of xanthophylls. For the above reasons, and with improved methods and equipment available, it was decided to reinvestigate the pigments of *Thoracophelia* and their metabolism in a more searching manner than was previously possible.

GENERAL DESCRIPTION AND HABITAT

Thoracophelia mucronata is a blood-red worm of some 0.5 to 2.0 mm. in diameter and from about 30 to 50 mm. or more in length when mature. Sexes are separate. Adults examined in May, July and September contained ripe eggs and sperm. The animals live buried in intertidal sand, where they may be dug in large and concentrated numbers from a fraction of an inch to a foot or more beneath the sand surface during low or medium tides. The worms swallow much sand, readily seen to constitute the gut-contents and feces of freshly collected specimens. When collected in their wet sand and placed on damp paper or in sea water, the worms crawl out of the sand in a few hours, clearing their bodies externally and partially internally of sand and other detritus, thus preparing themselves for analysis.

Sand grains in the gut of the worm and in its feces do not differ significantly in size or appearance from the grains of the general environment; minute algal cells and detrital matter may be observed on the surface of such grains. The worms' biology, food and migrations within the sand constitute the subject of a separate research project.

The bright red and purplish colors of the worms are due to the presence of hemoglobin, which is not enclosed in erythrocytes, but is free in colloidal solution. The shed blood of this species was not observed to clot on exposure to air. Oxyhemoglobin was identified by its two characteristic sharp absorption bands, centering at 576 and 540 mμ respectively. Reduction with sodium hydrosulphite gave the single broad band of hemoglobin at 565 mμ (Hartridge reversion spectroscope). A strong positive test for hemoglobin was obtained with the benzidine reaction.

CAROTENOIDS AND CHROMOLIPOIDS

Methods. The worms were usually permitted the minimum interval of an overnight period in which to emerge and clear their bodies largely of sand and detritus. And, since we wished to minimize both incomplete digestive changes and aberrant post-mortem bacterial operations, final procedures sometimes involved the tedious step of hand-picking the emerged survivors. After blotting on towels, masses of the animals were weighted and comminuted to a fine mush under acetone in a blender.

Carotenoids were analyzed in accordance with regular methods applied to animal tissues (cf. Fox and Pantin, 1941). Exhaustive acetone extracts of the ground tissues were pooled, transferred to petroleum ether by dilution and extraction, then partitioned into epiphasic fractions (remaining in the petroleum ether phase) and hypophasic portions (migrating into the 90% methanol layer below). Epiphasic fractions were treated separately with warm alcoholic NaOH and subsequently repartitioned. Petroleum ether solutions of the final epiphasic and hypophasic pigments were separately washed, dried with solid NaCl and chromatographed. Samples of whole pigment extracts or of various separated fractions were examined spectroscopically in CS_2 or in "Skellysolve B," a refined hydrocarbon boiling between 60° and 70° C.

Skellysolve solutions of the pigments were chromatographed on Tswett columns consisting of MgO mixed with Hyflo Supercel (2 : 1 by weight) overlaid by similar columns of $CaCO_3$—Supercel mixtures of like proportions. The only colored zones from either epiphasic or hypophasic pigments were formed in the MgO portions of such columns.

Carotenoid material was measured in mg. per 100 g. of sand (dry basis) or worms (wet), with the use of a Bausch and Lomb spectrophotometer, the absorption of light at 485 mμ by CS_2 solutions of pigment giving values in β-carotene equivalents. It is to be remembered, however, that recorded determinations may depart from true values, since we were dealing with lipoidally contaminated extracts rather than with solutions of pure carotenoids.

Results. Samples of sand, containing none of the worms, but from the same environment, were analyzed for pigments. The first 3.5 kg. sample yielded approximately 0.06 mg. of collective xanthophylls (in "lutein" units) per 100 g. of sand on the dry basis. The xanthophylls were accompanied by unmeasured amounts of carotenes, probable chlorophylls and greenish degradation products of the latter. A second sample of sand from the environment of a worm colony was given more thorough study. It weighed 4.1 kg. when dried, and contained 0.70% of organic matter (by combustion). Extracts of the fresh, undried sand yielded the following approximate quantities of biochromes, separated by partition and chromatography:

Epiphase	*In whole sand (mg. per 100 g.)*	*In organic matter of sand (mg. per 100 g.)*
"Chlorophyll"	0.041	5.85
"Phaeophytin"	0.016	2.30
β-carotene	0.002	0.28
α-carotene	0.001	0.14
Hypophase		
Green breakdown products (in "chlorophyll" units)	0.025	3.60
Total carotenoids (*i. e.*, five xanthophyllic fractions, preponderantly fucoxanthins, with chief maxima at 508–510 and 477 mμ in CS_2)	0.064	9.04

It is observed that β-carotene occurred in twice the concentration of the α-isomer, and that collective xanthophylls constituted 95.5% of the total carotenoids.

While actual amounts and ratios of the pigments in the sand vary with locality, depth and season, this general condition is characteristic of fresh, mixed marine detritus, and lends emphasis to the biochemically selective behavior of the worm, which selectively stores the β-carotene, eliminates the α-carotene, and apparently destroys the xanthophylls at least in part.

TABLE I. CAROTENOID AND CHROMOLIPOID MATERIALS IN *Thoracophelia mucronata*

Sample	*History*	*Weight* (*g.*)	*Total extract* (*mg. %*)	*Total epiphasic material* (*mg. %*)	*Total xanthophyllic material* (*mg. %*)	*Carotene (from hydrolysate)* (*mg. %*)
I	Recently collected, but some dead included. Crude extract hydrolyzed.	218.7	——	——	0.01 (with chromolipoids)	0.15
II	Recently collected but some dead included.	75.5	0.30	0.24	0.01 (with chromolipoids)	0.215
III	Blood only, from freshly collected worms.	12 ml.	0.11	——	none (only chromolipoids)	0.04
IV	Collected previous day. Only live worms used.	75.5	0.25	0.18	none (only chromolipoids)	0.11
V	Starved one week in clean gravel and running, filtered sea water at 20° C.; only surviving worms used.	35.6	0.315	0.24	0.015 (with chromolipoids)	0.12
VI	Collected, washed and immersed in acetone within 3 hours.*	57.9	0.50	0.36*	0.01 (with chromolipoids)	0.24

* This extract contained some α-carotene (*i. e.*, about 13% of the total carotenes, the other 87% being β-carotene) due to the presence of undigested food residues in the gut, since the animals were purposely analyzed while the gut was still filled with sand.

That extracellular chlorophylls and xanthophylls are largely or completely decomposed in the gut of the worm was indicated first by analyzing fresh worms within three hours of collecting them, when their alimentary tracts still contained considerable sand, and second by a preliminary analysis of about 11 g. of sand voided by previously washed and segregated worms.

It will be noted (Table I, Sample VI) that there was an increase in the average amount of total extracted yellow pigment and of β-carotene itself in the fresh worms. Furthermore, traces of α-carotene were encountered, doubtless derived from the intestinal contents of the fresh sample. But no significant quantities of xanthophylls were observed.

The concentrated extract of the fecal sand was yellow, rather than green, in contrast to extracts of the fresh sand. On a chromatogram it yielded small amounts of chlorophyll-degradation products, carotene in a concentration of the same order of magnitude as that encountered in fresh sand (*i. e.*, 0.004 mg. per 100 g. sand) and, on the $CaCO_3$ portion of the column, no xanthophylls, but some yellow pigment which may have been derived from xanthophyll. It was hypophasic, but gave no positive tests for carotenoid material either with $SbCl_3$ or with concentrated H_2SO_4; furthermore, its absorption spectrum in CS_2 (Beckman instrument) showed maxima displaced farther toward the violet than those of fucoxanthin, their centers lying at 441 and 474 mμ, with a slight inflection at 498 to 499 mμ. While it would thus seem that fucoxanthins were destroyed in the gut of the worm, the bare possibility has not been excluded that such carotenoids may be desorbed from the sand (while β-carotene is not) and discharged in fluid feces.

The worms' bodies yielded yellow pigments which were chiefly epiphasic. The only consistently encountered carotenoid was a persistently epiphasic, alkali-resisting carotene whose absorption maxima were found to agree with those of β-carotene from plant sources, *i. e.*, 427–430, 450, and 476 mμ in Skellysolve (Fig. 1). The identity of the *Thoracophelia* pigments fraction with β-carotene was further confirmed by the results of the three-tube mixed chromatogram method: (1) β-carotene; (2) worm carotene; (3) a mixture of these two; in each of the three Tswett columns of MgO there appeared a single, identical orange zone of adsorbed pigment. Traces of more readily adsorbed pigments usually appeared on primary chromatograms of the persistently epiphasic fraction. Such material probably represented artifactual isomers, colored degradation products or other contaminants. But no α-carotene, which accompanies the β-isomer in this and all other marine sediments and commonly appears as a yellow zone be-

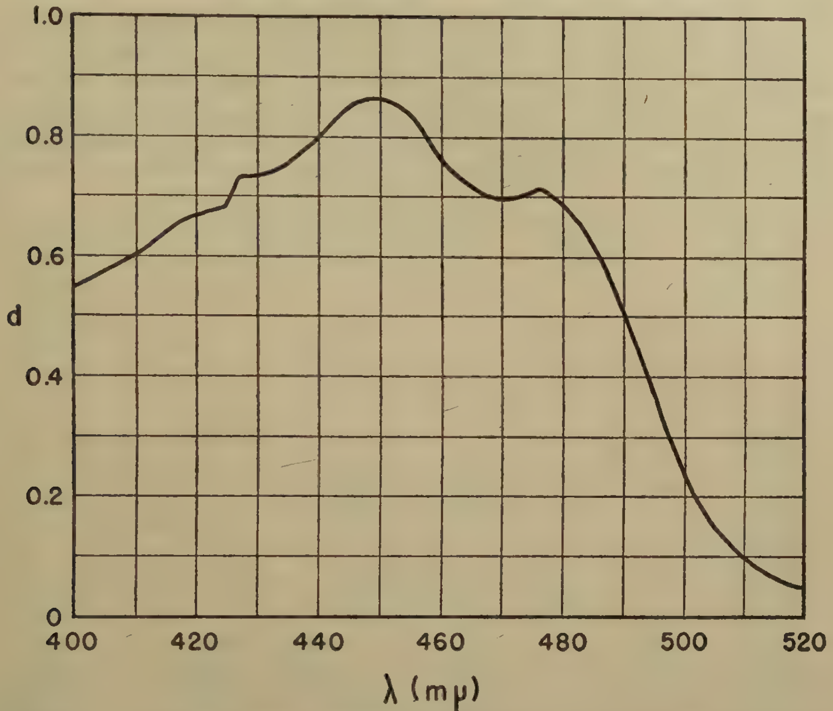


Figure 1. Absorption curve (Beckman) of β-carotene from *Thoracophelia*. Solvent Skellysolve B.

neath the orange band of the latter, was ever detected in extracts of the worm, save in traces when extracts were prepared from the fresh, actively feeding animals (Table I, Sample VI). The worm's assimilation of β-carotene, while excluding or destroying not only xanthophylls but even the isomeric α-carotene, all of which occur in the natural diet (Fox, *et al.*, 1944), seems remarkable. Quantitative spectrophotometric estimations of carotene in the worms' tissues are summarized in Table I.

The total yellow pigment values of 0.30, 0.25, 0.315, and 0.50 mg. %, measured as β-carotene equivalents on the raw extracts, agree rather well with that of 0.38 mg. % found earlier by a less refined, nonspectroscopic photometric method (Sumner and Fox, 1933).

Both epiphasic and hypophasic fractions yielded small amounts of yellow noncarotenoids which exhibited properties characteristic of the so-called chromolipoids or lipofuscines. Such pigments passed unadsorbed through $CaCO_3$, but could be chromatographed on MgO columns. More of this material was hypophasic than epiphasic. Most of the chromolipoid pigment which was initially epiphasic could be rendered hypophasic by treatment with alcoholic NaOH. A common property of the various fractions was the manifestation of blue fluorescence when dissolved in Skellysolve and illuminated with the beam of a carbon arc or with light through an ultraviolet filter. These

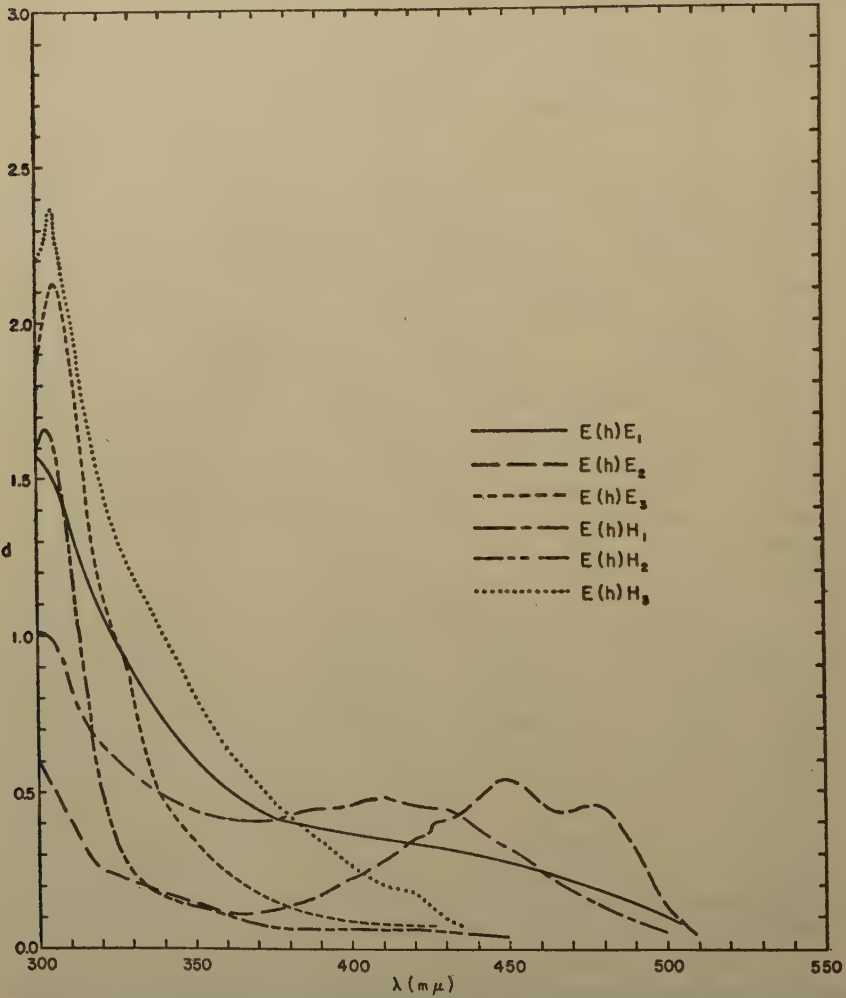


Figure 2. Absorption spectra (Beckman) of the persistently and initially epiphasic pigment (β-carotene and chromolipoids) from *Thoracophelia*. Solvent: Skellysolve B. Lettered fractions as in Table II.

chromolipoids were readily distinguishable from carotenoids by their fluorescent properties, by their absorption spectra, and by their failure to exhibit blue or green colors, showing instead brown, brown-yellow or brown-orange when treated with concentrated H_2SO_4. Chief among these distinguishing characters were the absorption spectra (Figs. 2, 3), which showed only slight absorption in the blue-green, blue

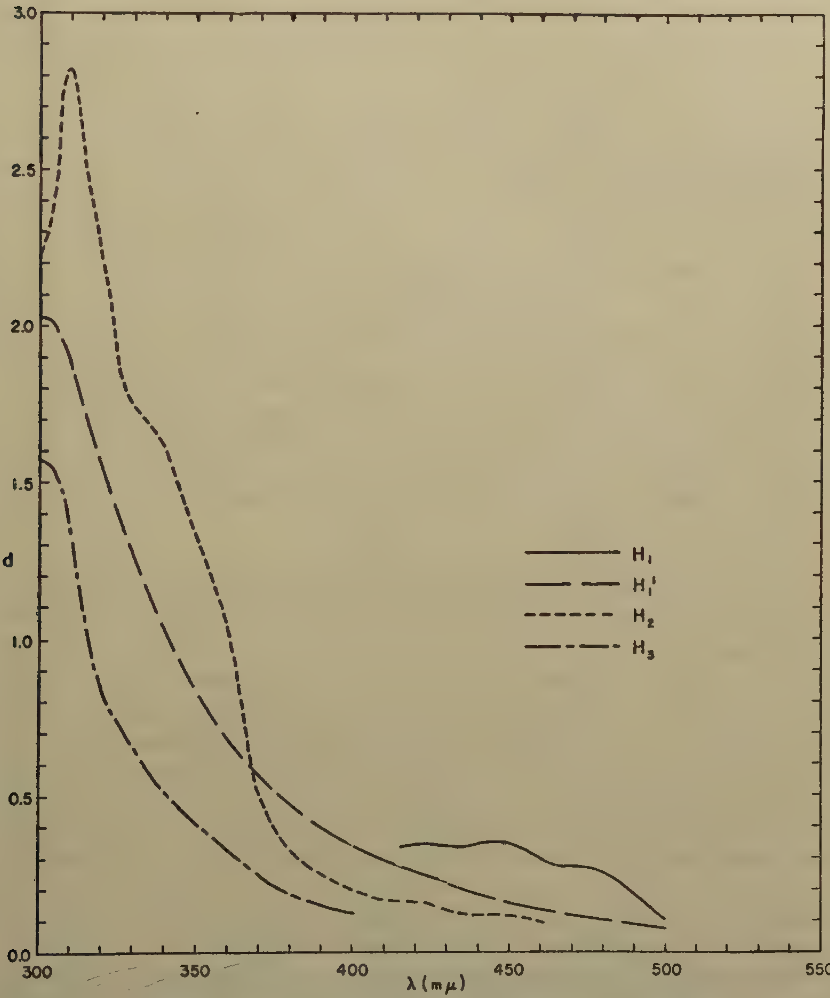


Figure 3. Absorption spectra (Beckman) of the hypophasic chromolipoids of *Thoracophelia.* Solvent: Skellysolve B. Lettered fractions as in Table II.

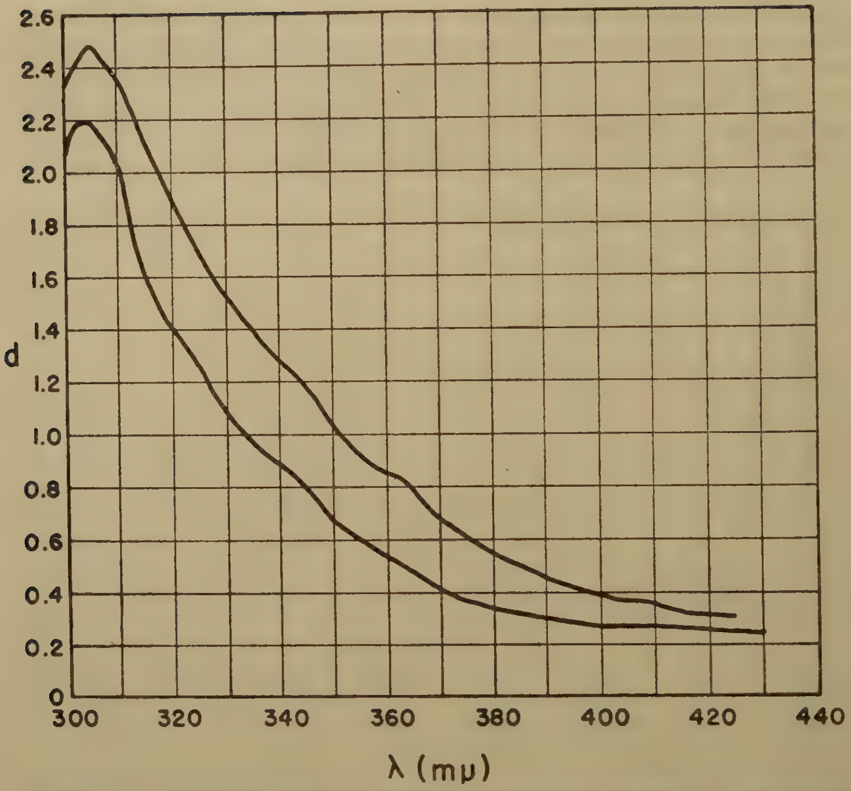


Figure 4. Comparative absorption spectra of a hypophasic chromolipoid (upper curve) and an apparently identical fraction derived by hydrolysis of an originally epiphasic, presumable ester (lower curve). Beckman instrument. Solvent: Skellysolve B.

and blue-violet, but which demonstrated steeply rising absorption in the violet, often with a sharp maximum at values from 300 to 310 mμ. Similar absorption curves of some initially hypophasic fractions, and of originally epiphasic ones rendered hypophasic by hydrolysis, suggest the presence in the tissues of a free and an esterified alcoholic component of some kind (Fig. 4). The absence of colored insoluble carotenoid salts at the interface of diluted alcohol-NaOH hydrolysates and petroleum ether indicated the absence, from either epiphasic or hypophasic extracts, of any acidogenic carotenoids commonly encountered in crustaceans, echinoderms, coelenterates and mollusks.

Treatment of the chromolipoid fractions with $SbCl_3$ in $CHCl_3$ solution failed to show the blue colors which characterize Vitamin A and numerous colored carotenoids, but resulted instead in pink to mauve colors. The absorption spectra, color of fluorescence, and hypophasic behavior of our chromolipoid fractions eliminate any identity with phytofluene, the widely occurring plant polyene hydrocarbon described by Zechmeister and Sandoval (1946). While the Liebermann-Burchard reaction demonstrated the presence of sterols in a number of the fractions, it is not believed that such compounds were responsible for the yellow color, since some yellow fractions seemed to contain no sterols, and since the better known sterols are colorless, absorbing light only in the ultraviolet range.

It is of interest to note, however, that the great majority of the chromolipoids were apparently neutral compounds. Two fractions only, H_1' and $E(h)E_1$, were not eluted by neutral acetone or acetone-Skellysolve systems, but could be removed from the MgO after acidification of the eluant with acetic acid. Therefore, these two portions may have had acidic properties, while the behavior of the others suggests higher alcohols.

Information relative to the lipofuscine fractions separated from sample V is summarized in Table II.

Four out of six lots of worms analyzed yielded traces of a hypophasic xanthophyll-like component or degradation product. This fraction departed from typical xanthophyllic behavior in that it passed through $CaCO_3$ Tswett columns unadsorbed. But it was adsorbed as a pink zone above the fluorescent noncarotenoid on the MgO portion of the column. Absorption maxima in CS_2 lay between 475–485 mμ and 505–510 mμ (Bausch and Lomb visual spectrophotometer) and in Skellysolve at 422–425, 446 and 472 mμ (Beckman photoelectric instrument). Such values are reminiscent of those of lutein (420, 446.5, 476 mμ in petroleum ether or ethanol) and taraxanthin (417, 443, 472 mμ in ethanol), but are not identical with either set of values. No such material was recovered from hydrolysates of epiphasic fractions. Because of its inconsistent occurrence, the minute traces in which it was found when present, its atypical chromatographic behavior and the fact that no corresponding esters were ever found, we believe that this pigment may reflect either the decomposition of carotene or the presence of undegenerated food-residues in the intestinal tract of the animal. The inadsorbability of any yellow *Thoracophelia* pigments on $CaCO_3$ had been observed in the earlier work of Sumner and Fox (1935).

In its prevention of xanthophylls from entering the system, whether by not permitting diffusion through the intestinal mucosa or by destruction of such pigments in the lumen of the gut, *Thoracophelia* resembles some higher herbivores or omnivores rather than many of the lower forms. Cattle are known to absorb a part of the dietary

TABLE II. PHYSICAL AND CHEMICAL PROPERTIES OF CHROMATOGRAPHIC PIGMENT FRACTIONS FROM *Thoracophelia*

Fraction (*color on* MgO *column*)	*Blue fluorescence*	*Absorption max.* (*m*μ)	$SbCl_3$ *reaction*	*Liebermann-Burchard reaction for sterols*		*Acid treatment*	
				H_2SO_4 *layer*	$CHCl_3$ *layer*	H_2SO_4 *layer*	$CHCl_3$ *layer*
H_1 (pink)	+	<300	pink	purple	green	....	....
H_1' (pink)	+	ca. 300; 423; 446; 472	murky	brick-red	brown	....	....
H_2 (purplish-gray)	+	310 (sharp)	red-orange	purple	blue-green	....	....
H_3 (no colored zone)	+	ca. 300	pink (±)	faint purple	pale yellow	....	....
$E(h)E_1$ (orange)	+	<300	pale pink	purple	blue-green	....	....
$E(h)E_2$ (orange)	+ (impurity)	(427); 450; 477 (β-carotene)	blue	green or purple	colorless to blue	green	brown (impurities)
$E(h)E_3$ (diffuse yellow)	+	306 (sharp)	pink (±)	purple	brown	....	....
$E(h)H_1$ (reddish)	(+)	ca. 300; (390); (410); (428)	–	–	–	pale brown	bleached
$E(h)H_2$ (yellow)	(+)	302.5 (sharp)	–	–	–	pale yellow-brown	bleached
$E(h)H_3$ (yellow filtrate)	(+)	305 (sharp)	pink	deep blue	green	orange	brown-orange

H = hypophasic pigment
E = epiphasic pigment
(*h*) = hydrolytic operation
E(*h*)*E* = persistently epiphasic pigment
E(*h*)*H* = originally epiphasic pigment rendered hypophasic by hydrolysis.

Subscript numbers refer to chromatographic position, beginning at top.

H_1' = fraction of *H*, elutable with acetic acid; similarly with $E(h)E_1$.
(+) = very faint positive
± = questionable
– = no reaction detected.

Maxima recorded in parentheses indicate very weak ones or points of inflection.

carotenes but very little xanthophyllic compounds (Palmer and Eckles, 1914; Goodwin, Dewar and Gregory, 1946). Sheep store mere traces of carotene and Vitamin A in their blood and other tissues (Palmer, 1916; Paulsen, Hilmoe and Moxon, 1944; Peirce, 1945) while excreting in their feces 70% of ingested carotenes and 80% of the xanthophylls consumed (Rogozinski, 1937). Vitamin A introduced by abdominal cannula into the digestive tract of sheep and goats produces a rise of Vitamin A levels in both systemic and portal blood, but colored carotenoids so introduced are not reflected in the blood in this way (Goodwin, *et al.*, 1946). It is believed that sheep must destroy xanthophylls in the digestive tract (Rogozinski, 1937). The same view is held regarding the horse, which assimilates only carotenes from green food, but excretes 70–80% of the dietary carotenes and 60% of the xanthophylls ingested. Since none of the latter are assimilated as such, they must be partially destroyed in the gut (Zechmeister and Tuzson, 1934; Zechmeister, Tuzson and Ernst, 1935). Hove (1943) found that water-extracts of minced rat stomachs destroyed carotene rapidly in vitro, in the presence of methyl linolate. He suggests the possible presence of carotene oxidase or lipoxidases. Sherman (1947) has shown that gastro-intestinal destruction of dietary carotene in the rat may be decreased by the administration of lutein. He suggests that the effect of the lutein may be due to its preferential destruction by a nonspecific enzyme which may be responsible for the degradation of carotenoids generally. The cod, *Gadus callarias*, and the lobster, *Homarus americanus*, are believed to slowly convert carotene into Vitamin A, as do mammals and birds (Neilands, 1947; Mattson, *et al.*, 1947; Wiese, *et al.*, 1947).

No destruction of xanthophylls from sand were demonstrable in preliminary experiments involving the incubation of such material (adsorbed in $CaCO_3$) with pastes of ground *Thoracophelia*. Partially starved worms (Table I, Sample V) exhibited no significant changes in carotene content, but did show slightly elevated quantities of the yellow noncarotenoid chromolipoid material (absorption at 305 mμ about 3 : 1).

In its assimilation of β-carotene without the accompanying α-isomer, *Thoracophelia* seems to be unique. This finding is contrary to an earlier provisional conclusion that the carotene in this species was preponderantly the α-isomer (Fox, 1936).

The reinvestigation of carotenoids in *Thoracophelia* lends emphasis to a provisional conclusion reached earlier by Sumner and Fox (1935) regarding the probable conversion of carotene into xanthophyll by the Pacific killifish, *Fundulus parvipinnis*, which stores only esterified members of the latter class. Fishes maintained for three months upon carotenoid-free food neither lost nor gained xanthophyll, while parallel groups fed upon xanthophyll-rich flesh, and those kept on a diet of *Thoracophelia* underwent increases in total amounts of stored xanthophyll. Chemical examinations made at the time of those experiments and subsequently (Fox, 1936) revealed none of the yellow noncarotenoid chromolipoids of *Thoracophelia* in the fishes which ate the worms, since hydrolyzed extracts of the fish pigments were quantitatively adsorbed upon $CaCO_3$ columns and exhibited well defined absorption spectra.

Nor could the increase of xanthophyll stored by the worm-fed fishes have been logically assigned to the traces of "xanthophyll" sometimes encountered in the worms, for the average amount of residual "xanthophyll" revealed by the present survey is, according to Table I, about 0.031 mg. from the collective 475.2 g. of worms, or some 0.006 mg. % on the average. Sumner and Fox (1935) fed their worm-eating lot of *Fundulus* about 2000 g. of this food, corresponding to roughly 0.12 mg. of "xanthophylls" (and 6 mg. of carotene) over the three month experimental period. During that interval the growing consumers, while undergoing a decrease in *concentration* of stored xanthophylls, nevertheless achieved an actual collective augmentation of approximately 0.20 mg. of this class of carotenoid. Even quantitative absorption of all available dietary xanthophyll (which is not known to occur in any animal, and certainly was not achieved by the xanthophyll-fed lot in the experiment cited) would have fallen 40% short of supplying the observed increment. Since synthesis of xanthophylls from colorless precursors did not appear to take place in the fishes receiving the carotenoid-free diet of halibut flesh, we are still left with the conclusion that *Fundulus* must be able to convert nonxanthophyllic into xanthophyllic carotenoids. Furthermore, the only carotenoid which the fishes could have received from *Thoracophelia* was β-carotene.

COMPOSITION, FOOD AND METABOLISM

The finding in the worm's gut of little else of a microscopically visible nature save sand grains led to a consideration of the basic food of the animal.

The gut wall consists of deep and multiple longitudinal folds throughout its anterior portion and is provided throughout its length with rapidly sweeping cilia upon its inner surfaces. Peristalsis is marked. These factors, added to the rich blood supply in the intestinal walls, provide for intimate contact with the ingested sand grains in the region of digestion and absorption.

Beyond question, the organic material present in the sand exists largely as adsorbed colloidal micelles and other fine particles. This adsorptive property of sand, which had previously been chemically cleaned, was demonstrated in the laboratory by gentle agitation with colloidal material prepared by the fine comminution of marine plants, followed by centrifugal removal of the coarser particles. Extraction of the treated and subsequently water-rinsed sand with acetone yielded green and yellow plant pigments. Egg albumin, the colloidal dye congo red, and the crystalloid organic dyes night blue, methyl violet and methylene blue were also observed to be readily and firmly adsorbed to sand. Congo red was adsorbed more readily to raw, untreated sand, wherefrom it could be removed incompletely and with difficulty by vigorous shaking, while the other dyes became adsorbed in greater quantities to the previously cleaned sand. These observations suggested that the marine organic colloids already adsorbed upon the raw sand must inhibit somewhat the additional adsorption of crystalloidal dyes, while increasing that of the electrically charged colloid dye congo red.

Use was made of the firm adsorbability of the nontoxic dye, methylene blue, upon sand to demonstrate the approximate rate of sand-ingestion by the worms. The dyed sand was washed with sea water until no further dye was thus desorbed and washed out, but still leaving the sand a deep blue color. Freshly collected *Thoracophelia* were then buried in the damp blue sand contained in earthenware plant-pots, standing partially immersed in sea water. Subsequent examination disclosed that numerous specimens had evacuated the raw sand from the gut and had replaced this with the dyed sand within 15-minute intervals. The dyed sand was in turn readily replaced by normal sand when the worms were given the opportunity. They seem to swallow either substrate equally readily.

The proportions of water, ash and total organic matter in fresh worms were determined, as well as the relative quantities of sand which had been ingested (all as listed in Table III). The values for sand were determined by treatment of the ash with dilute HCl, followed by reweighing the dried insoluble residue. The results obtained are doubtless conservative, since carbonates and iron sulfide are soluble in dilute acids and therefore would have been removed by the treatment accorded the ash.

Some rough calculations may be useful in suggesting the quantities of sand and organic matter adsorbed thereto, which the worms may ingest and utilize in order to build up or maintain the organic material of their tissues.

TABLE III. Composition of Fresh Worms Corrected for Quantities of Ingested Sand (Silica)

	Sample 1	*Sample 2*	*Average*
Wet weight, g.	40.97	37.33	.....
Dry matter, *corrected for silica*, %	19.82	19.70	19.76
Water (by difference), %	80.18	80.30	80.24
Ash, *corrected for silica*, %	2.02	2.02	2.02
Organic matter (by difference), %	17.80	17.68	17.74
Silica in gut (g. per 100 g. of fresh worm tissue)	5.84	5.29	5.56

The whole worms contain close to 18% organic matter and swallow some 6% of their own weight of sand. This sand may contain about 1% of organic matter, *i. e.*, by taking a rough average between our figure of 0.7% and the value of 0.1% of organic nitrogen (given by Revelle and Shepard, 1939), the latter value being multiplied by a factor of 14 to 18 to give the total organic matter (Trask, 1939). Thus, if 100 g. of the worms fill their alimentary tracts each quarter hour with 6 g. of sand, this represents a cycle of more than 210 kg. of sand per annum, or of somewhat more than 2000 g. of organic matter adsorbed to the sand. Grown specimens weigh about 40 mg. each and therefore ingest on the average some 9 mg. of sand per hour, 0.23 g. per day, or about 84 g. per year. Adult animals would appear to operate at a biochemical efficiency of 18/2000 x 100 or 0.9% during a year's time. And it is worth re-emphasizing that the worm, while assimilating none of the relatively plentiful xanthophylls or α-carotene from the organic matter in the sand, contrives to concentrate the β-carotene by some five-fold, if we take 0.28 mg. of β-carotene as its concentration in the organic matter of the sand and 100/18 x 0.25 or 1.4 mg. as its approximate concentration in the worm's organic matter.

Some approximate values are given in tabular form below, to indicate the relative magnitude of local operations of the beach-worm. Populations are based upon counts of samples, and values for the density of sand, the rate of its ingestion, etc., are based upon measurements of samples.

Volume of damp sand in worm-bed, 1 mile long, 10 ft. wide, 1 ft. thick	5.28×10^4 ft.^{3}
Weight of damp sand in this volume (density ca 131 lbs/ft.^{3})	3.46×10^3 tons
Weight of actual sand (subtracting ca 18.5% water)	2.82×10^3 tons
Total number of worms in worm-bed (ca 3×10^3/ft.^{3})	1.58×10^8
Total weight of worms in worm-bed (ca 0.04 g. each)	7.0 tons
Weight of sand cycled annually by worms (ca 84 g. each)	1.46×10^4 tons
Ratio of sand cycled annually by worms to sand in worm-bed	$\frac{1.46 \times 10^4}{2.82 \times 10^3} = 5.2$

(equivalent to all the sand in the worm-bed about once every 10 weeks)

Organic matter in sand cycled annually by worms (ca 1%)	1.46×10^2 tons
Organic material in worms in the colony (ca 20%)	1.4 tons
Relative efficiency of worms in maintaining a standing population, assuming an annual life-cycle	$\frac{1.4 \times 10^2}{1.46 \times 10^2} = 0.96\%$
Quantity of carotene in sand of worm-bed (ca $2.8 \times 10^{-4}\%$)	16 lbs.
Quantity of carotene in worms of the colony (ca $2.5 \times 10^{-4}\%$)	0.035 lbs.
Relative efficiency of carotene assimilation, assuming an annual life cycle	$\frac{0.035 \times 10^2}{16 \times 5.2} = 0.042\%$
Volume of sand in wedge of beach, 1 mi. long, 200 ft. wide, 2.5 ft. high	1.32×10^6 ft.^{3}
Weight of dry sand in wedge	7.05×10^4 tons
Fraction of sand on beach cycled annually by worms	$\frac{1.46 \times 10^4 \times 10^2}{7.05 \times 10^4} = 21\%$

SUMMARY

1. The blood-red, hemoglobin-containing, sand-eating marine polychaete worm, *Thoracophelia mucronata*, carries in its blood and other tissues appreciable quantities of β-carotene.

2. No α-carotene or other carotenoids could be detected in the worms' tissues, save for inconsistent traces of an unesterified, hypophasic, xanthophyll-like pigment or colored degradation product which was unadsorbed by $CaCO_3$.

3. Small quantities of yellow, noncarotenoid blue-fluorescent chromolipoids are extractable. Some of these are hypophasic initially, while others are rendered so by hydrolysis. These chromolipoids exhibit low light-absorption in the green, blue and blue-violet regions, but single sharp maxima in the ultraviolet at 310, 306, 305, 302.5, 300 mμ or below.

4. Although no carotenoid oxidase was demonstrated, it is suspected that the remarkable exclusion of xanthophylls from its system by *Thoracophelia* may be allied to the presence of lipoxidases, which must participate in the genesis of the animal's yellow chromolipoids.

5. *Thoracophelia* probably derives its subsistence from the colloidal organic matter adsorbed to the sand. Consuming some 24% of their own weight of sand per hour, 100 g. of the animals must ingest some 210 kg. of sand annually, representing about 2000 g. of organic matter. Since each adult worm weighs about 40 mg., this is equivalent to about 84 g. of sand (containing .84 g. organic matter) per worm per year.

REFERENCES

Drumm, P. J. and W. F. O'Connor
1940. Pigments of sponges. Nature, Lond., *145;* 425.

Drumm, P. J., W. F. O'Connor and L. P. Renouf
1945. The pigments of sponges. 1. The lipid pigments of the sponge *Hymeniacidon sanguineum* (Grant). Biochem. J., Cambr., *39:* 208–210.

Fox. D. L.
1936. Further studies of the carotenoids of two Pacific marine fishes, Fundulus parvipinnis and Hypsypops rubicunda, and of a marine annelid, Thoracophelia sp. Proc. nat. Acad. Sci., *22:* 50–54.

Fox, D. L. and C. F. A. Pantin
1941. The colours of the plumose anemone, *Metridium senile* (L). Philos. Trans., (B) *230:* 415–450.

Fox, D. L. and B. T. Scheer
1941. Comparative studies of the pigments of some Pacific coast echinoderms. Biol. Bull. Woods Hole, *80:* 441–455.

Fox, D. L., D. M. Updegraff and G. D. Novelli
1944. Carotenoid pigments in the ocean floor. Arch. Biochem., *5:* 1–23.

Goodwin, T. W., A. D. Dewar and R. A. Gregory
1946. The transport of absorbed carotene in herbivora. Biochem. J., *40*, Proc. Biochem. Soc., Cambr., lx–lxi.

Hove, E. L.
1943. The in-vitro destruction of carotene by water extracts of minced rat stomachs in the presence of methyl linolate. Science, *98*(2550): 433–434.

Lederer, Edgar
1938. Recherches sur les carotenoides des animaux inferieurs et des cryptogames. Lons-le-Saunier, Paris.

Mattson, F. H., J. W. Mehl and H. J. Deuel, Jr.
1947. Studies on carotenoid metabolism. VII. The site of conversion of carotene to Vitamin A in the rat. Arch. Biochem., *15:* 65–73.

Neilands, J. B.
1947. The conversion of carotene to Vitamin A in the fish. Arch. Biochem., *13:* 415–419.

Palmer, L. S.
1916. The physiological relation of plant carotinoids to the carotinoids of the cow, horse, sheep, goat, pig, and hen. J. biol. Chem., *27:* 27–32.

Palmer, L. S. and C. H. Eckles
1914. Carotin—The principal natural yellow pigment of milk fat: its relations to plant carotin and the carotin of the body fat, corpus luteum and blood serum.

I. The chemical and physiological relation of the pigments of milk fat to the carotin and xanthophylls of green plants. J. biol. Chem., *17:* 191–210.

II. The pigments of the body fat, corpus luteum and skin secretions of the cow. J. biol. Chem., *17:* 211–221.

III. The yellow lipochrome of blood serum. J. biol. Chem., *17:* 223–236.

IV. The fate of carotin and xanthophylls during digestion. J. biol. Chem., *17:* 237–243.

V. The pigments of human milk fat. J. biol. Chem., *17:* 245–249.

PAULSEN, T. M., R. J. HILMOE AND A. L. MOXON

1944. The carotene content of sheep blood. Proc. S. Dak. Acad. Sci., *24:* 81–84. (Chem. Abstr., *39:* 4936 [1945]).

PEIRCE, A. W.

1945. The effect of intake of carotene on the general health and on the concentration of carotene and of Vitaimin A in the blood and liver of sheep. Aust. J. exp. Biol. med. Sci., *23:* 295–303.

REVELLE, ROGER AND F. P. SHEPARD

1939. Sediments off the California Coast; in, Recent Marine Sediments; pp. 245–282. Amer. Ass. Petrol. Geol., Tulsa, Okla. 736 pp.

ROGOZINSKI, F.

1937. Carotenoids and chlorophyll in the digestion of the ruminant. Bull. int. Acad. Polon. Sci., Cl. Sci. Math. Nat. (B) *11:* 183–193 (Chem. Abstr., *32:* 1760 [1938]).

SCHEER, B. T.

1940. Some features of the metabolism of the carotenoid pigments in the California sea mussel (Mytilus californianus). J. biol. Chem., *136:* 275–299.

SHERMAN, W. C.

1947. Relative gastro-intestinal stability of carotene and Vitamin A and protective effect of xanthophyll. Proc. Soc. exp. Biol., N. Y., *65:* 207–210.

SUMNER, F. B. AND D. L. FOX

1933. A study of variations in the amount of yellow pigment (xanthophyll) in certain fishes, and of the possible effects upon this of colored backgrounds. J. exp. Zool., *66:* 263–301.

1935. Studies of carotenoid pigments in fishes. III. The effects of ingested carotenoids upon the xanthophyll content of Fundulus parvipinnis. Proc. nat. Acad. Sci., *21:* 330–340.

TRASK, P. D.

1939. Organic content of recent marine sediments; in, Recent Marine Sediments, pp. 428–453. Amer. Ass. Petrol. Geol., Tulsa, Okla. 736 pp.

WIESE, C. E., J. W. MEHL AND H. J. DEUEL, JR.

1947. Studies on carotenoid metabolism. VIII. The *in vitro* conversion of carotene to Vitamin A in the intestine of the rat. Arch. Biochem., *15:* 75–79.

ZECHMEISTER, L. AND A. SANDOVAL

1946. Phytofluene. J. Amer. chem. Soc., *68:* 197–201.

ZECHMEISTER, L. AND P. TUZSON

1934. Beitrag zum Lipochrom-Stoffwechsel des Pferdes. Z. physiol. Chem. (Hoppe-Seyl. Z.), *226:* 255–257.

ZECHMEISTER, L., P. TUZSON AND E. ERNST

1935. Selective accumulation of lipochrome. Nature, Lond., *135:* 1039.

NUTRITION, ENVIRONMENTAL CONDITIONS, AND GROWTH OF MARINE BIVALVE MOLLUSKS[1]

By

WESLEY R. COE

Osborn Zoological Laboratory, Yale University,
New Haven, Conn.
and
Scripps Institution of Oceanography, University of California
La Jolla, California

Recent investigations on the nutrition of the ciliary-feeding bivalves indicate that, although many varieties of organic and inorganic substances are ingested, most of the actual food is assimilated by the intracellular digestion of very small cells and particles. This is accomplished mainly by the phagocytic cells of the digestive organ (digestive diverticula, or "liver"), but it is aided to a small extent by migratory phagocytic cells in the lumen of the alimentary canal. The greater proportion of the ingested material consists of particles too large for this form of digestion, and the alimentary canal is deficient in most of those enzymatic extracellular secretions which are responsible for the primary digestive processes in many of the other invertebrates.

The principal extracellular secretions in the alimentary canal of these mollusks are amylase and glycogenase, secreted by the cells of the style-sac and incorporated in the style. These enzymes are set free in the stomach as the end of the style is dissolved, and they are then available for the digestion of the free starch and glycogen which may be contained in the food.

The gastric juice also contains a small amount of cellulase of the same origin, but this is evidently inadequate for the complete digestion of dinoflagellates and other phytoplankton with cellulose walls, which frequently constitute the greater part of the stomach contents. The nutritive substances contained in these organisms become available as nutrition for the bivalves only after death and disintegration into particulate materials of minute size. Only to a small extent does this occur during the short time that the ingested cells remain in the alimentary canal. Most of the living dinoflagellates pass through the canal without apparent change.

Because the stomach usually contains such a large proportion of dinoflagellates and relatively large diatoms, it was formerly assumed that these constituents of the phytoplankton contribute, as living cells, the greater part of the food material. However, Yonge (1926, 1935, 1946) has proved that only the smallest diatoms and similarly small organisms and particles can be assimilated by the oyster (*Ostrea lurida*). He found that the protease which would be required for the digestion of naked protoplasm or proteins could be detected in the lumen of the canal only in very small amount. Nor is there any considerable quantity of free lipase for the digestion of lipoids. The traces of these enzymes in the juices of the canal were ascribed to burst or cytolyzed phagocytes.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 376.

In his studies on the razor clam (*Ensis siliqua*), Graham (1931) likewise concluded that digestion in that species is mainly intracellular. However, he found small amounts of both amylytic and lipolytic enzymes in the midgut, with traces of the same enzymes in the rectum. He agreed with Yonge that they were due to the phagocytes present and that they were not to be considered as extracellular secretions. The style provides a strong extracellular amylase and an oxidase but no protease, lipase or invertase. Within the cells of the digestive organ (digestive diverticula) he demonstrated protease, lipase, diastase and glycogenase.

Yonge's conclusions have been supported also by the studies of Fox and Marks (1936), Fox and Coe (1943), and Coe and Fox (1942, 1944) on the sea-mussel (*Mytilus californianus*), as well as by those of Coe (1945) on the bay-mussel (*Mytilus edulis diegensis*) and Coe (1947) on the Pismo clam (*Tivela stultorum*). In all of these species, and in others less fully investigated, the gastric juice contains much amylase and glycogenase but only a small amount of cellulase. The rest of the nutrition is dependent upon the intracellular digestion of minute cells and particles within the cells of the digestive organ, supplemented by migratory phagocytes in the lumen of the alimentary canal. These cells contain within themselves the enzymes necessary for the digestion of carbohydrates, fats and proteins.

In addition to the indigestible remains of the phagocytized materials, the cells of the digestive organ excrete into the alimentary canal many of their characteristic brown granules. These are believed to be excretions and the remains of chlorophyll digestion, since all or most of them are discharged in the feces. It seems quite possible, however, that some of the digestive enzymes contained in the phagocytic cells may be secreted into the canal simultaneously, as Mansour and others maintain.

In the most primitive pelecypods (Protobranchia), digestion is reported by Yonge (1939) as being exclusively intracellular and within the cells of the digestive organ. No amylase is secreted into the stomach and there are no phagocytic cells in the lumen of the alimentary canal.

The investigations of Mansour (1946, 1946a, 1946b) and of Mansour-Bek (1946, 1946a) have led them to decide that the conclusions of the above mentioned writers are not applicable to all species of pelecypods. They report that the filtered, phagocyte-free stomach juices of *Tridacna* and *Pinctada*, and presumably of other genera, do contain proteolytic, lipolytic and carbohydrate-splitting enzymes, whereby even relatively large multicellular zooplankton can be digested. The conclusion was reached that these enzymes are set free in the canal by the fragmentation and cytolysis of ends of the phagocytic cells of the digestive organ, which are well known to contain all the enzymes mentioned. This is confirmed for *Unio* by Mansour and Zaki (1946).

A similar sloughing off of the distal portions of the phagocytic cells of the digestive organ has been reported in gastropods. Carriker (1946) and Carriker and Bilstad (1946), for example, found that in the snail *Lymnaea* the entire distal ends of some of the digestive cells are cast off in the final phase of secretion and pass into the lumens of the follicles "as an apocrine type of secretion." Nevertheless, no evidence of a protease could be detected in the gut fluids. The cells of the digestive organ here, as in the pelecypods, "function in secretion, assimilation, intracellular digestion and excretion," but in *Lymnaea* "assimilation also occurs in the pylorus and absorption possibly in the intestine." It is evident that extracellular digestion in the lumen of the canal is of minor importance.

Other gastropods, as *Crepidula* and *Vermetus*, resemble the pelecypods in their method of feeding by ciliary action and the secretion of mucus. They also resemble the pelecypods in having the alimentary canal provided with a crystalline style which supplies enzymes for the extracellular digestion of carbohydrates. The other nutrients, in the form of small particles only, are presumably digested intracellularly within the cells of the digestive organ as in the pelecypods (Yonge, 1932).

NUTRITION AND GROWTH

The rates of growth of the sea-mussel, as correlated with temperature and with the abundance of dinoflagellates, diatoms and bacteria in the water, were studied experimentally by Coe and Fox (1942, 1944) and by Fox and Coe (1943). In these experiments an average of about 100 individuals of carefully measured sizes were placed in separate compartments of wire-screened boxes and suspended below the low-water level in the vicinity of the natural mussel populations on the pier at the Scripps Institution of Oceanography. The mussels were measured at monthly intervals. In order that the boxes should always contain individuals of the same average size and age, the oldest group was discarded and replaced by young individuals after the former reached a length of about 80 mm., when growth in length decreases. The stomach contents were examined at frequent intervals. These experiments were continued for four years (1940 through 1943).

Similar studies were made by Coe (1945) on the bay-mussel, which also formed large colonies on the pier.

The nutrition and rates of growth of the Pismo clam were investigated by Coe (1947) by measuring at monthly intervals for two years (1945–1947) marked individuals living in the sand among the natural clam beds on the beach near the Scripps Institution. Large samples of the natural population were also measured at the same intervals and returned to the same positions on the beach after the measurements had been taken. Since these clams remained in their normal environment, the observed growth rates were presumably representative of the population for the two years of the experiment at the locality where the observations were made.

TABLE I. AVERAGE INCREASE IN LENGTH OF THREE SPECIES OF BIVALVE MOLLUSKS AS COMPARED WITH TEMPERATURE AND INCIDENT SOLAR RADIATION AND WITH ABUNDANCE OF DINOFLAGELLATES AND DIATOMS IN THE SAME LOCALITY DURING EACH MONTH. EACH OF THESE AVERAGES COVERS TWO YEARS. SOLAR RADIATION EXPRESSED IN TERMS OF g. cal. cm²/day.

	Av. Temp. (°C.)	*Solar radiation*	*Phytoplankton (Cells per liter)*		*Pismo clam Av. age 18 months* (mm.)	*Sea-mussel Av. age 6 months* (mm.)	*Bay-mussel Av. age 6 months* (mm.)
			Dinoflagellates	*Diatoms*			
Oct.	18.4	331	23,000	18,000	2.1	5.0	6.4
Nov.	15.6	303	10,000	4,000	1.7	4.3	5.9
Dec.	14.4	214	5,000	2,000	1.2	3.5	5.1
Jan.	13.2	246	2,000	1,000	0.9	3.2	3.8
Feb.	13.3	309	3,000	27,000	0.7	3.1	3.2
Mar.	14.2	461	12,000	20,000	1.2	4.1	4.2
Apr.	16.0	440	23,000	22,000	2.0	5.2	5.3
May	17.8	566	53,000	50,000	2.3	5.8	6.3
June	19.4	524	108,000	42,000	2.4	5.6	6.7
July	19.8	561	98,000	50,000	2.3	5.5	6.5
Aug.	20.6	508	20,000	107,000	2.0	5.3	6.3
Sept.	19.2	430	28,000	42,000	2.3	5.5	6.6
Average increase per year					21.1	56.1	66.3

RATES OF GROWTH

As indicated in Table I and Fig. 1, both the bay-mussel and the clam showed a maximum rate of growth in June and a minimum in January and February, the months of average lowest temperature. For the sea-mussel the maximum was in May and the minimum from December to February. All of the three species showed a slight decrease in

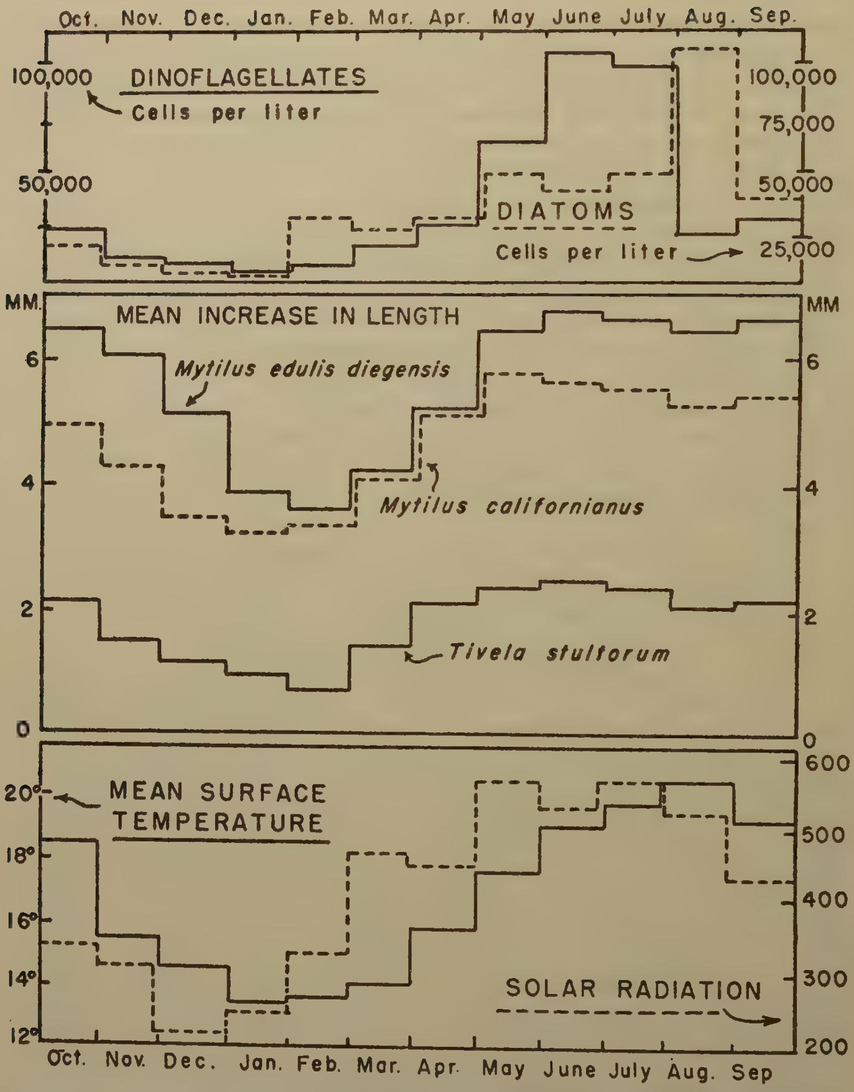


Figure 1. Correlation between temperature, solar radiation, abundance of dinoflagellates and diatoms and mean monthly increase in length of three species of bivalve mollusks under natural conditions at the Scripps Institution of Oceanography. The measurements indicated for each species represent the averages for two years. The data for the clam are based on measurements made in 1945–47, while the other data are taken from the records of the preceding years. The growth rates are more nearly parallel with the abundance of dinoflagellates than with the other environmental conditions. The incident solar radiation is shown in terms of g. cal. cm³ per day.

August, the month of highest temperature, followed by a rise in September.

Fox and Coe (1943) found that the average length of the sea-mussels in the experimental boxes at the age of one year was 56.1 mm., the maximum length observed at that age being 86 mm. There was a gain in length of about 45 mm. during the second year and 25 mm. during the third year. The dry weights of the body tissues at one, two and three years of age averaged 2, 5.6, and 8 g. respectively, and of the shell at the same ages 23.5, 70, and 104.59 g., respectively. It will be observed that while the increase in length during the second year was little more than half the increment of the first year, the volume, weight of tissues, and weight of shell was more than doubled.

The bay-mussel grew more rapidly than the sea-mussel under the same conditions during its first year but more slowly during its second year. A mean length of about 66.3 mm., with an observed maximum of 82 mm., was attained at the age of one year, and a maximum of about 93 mm. at the age of two years. Thereafter the rate of growth was very slow; few individuals in the local population exceed 105 mm.

The Pismo clam grows much more slowly than either species of mussel, the maximum observed length at one year of age being only 31 mm. and the average 21.1 mm. This rate shows but little diminution during the next two years, but decreases thereafter. In the local population a maximum length of 150 mm. with a shell weight of 600 g. at the age of ten to twelve years has been observed, but larger individuals have been reported from other parts of the coast. The growth rates of this species, as indicated by large collections of the population for each month of the two years 1945–47, as well as by marked individuals, are shown in Fig. 2. On some beaches on the California Coast, particularly those contaminated with sewage, the growth rates are greater and in other localities less than here recorded for La Jolla.

Sources of Nutrition

At all seasons of the year the greater part of the stomach contents of all the ciliary-feeding bivalves investigated consists of detritus and phytoplankton, particularly dinoflagellates and diatoms, together with much inert material, such as particles of sand, tiny fragments of mollusk shells and small pieces of algae. Included also, but in much smaller amount, are the more minute particles and organisms mentioned below as forming the principal part of the nutrition. There is little or no selection from among these particles except for size, although spiny objects are generally excluded.

The water about the mussel and clam beds at the Scripps Institution of Oceanography is reported to contain an average of about .005 g. of

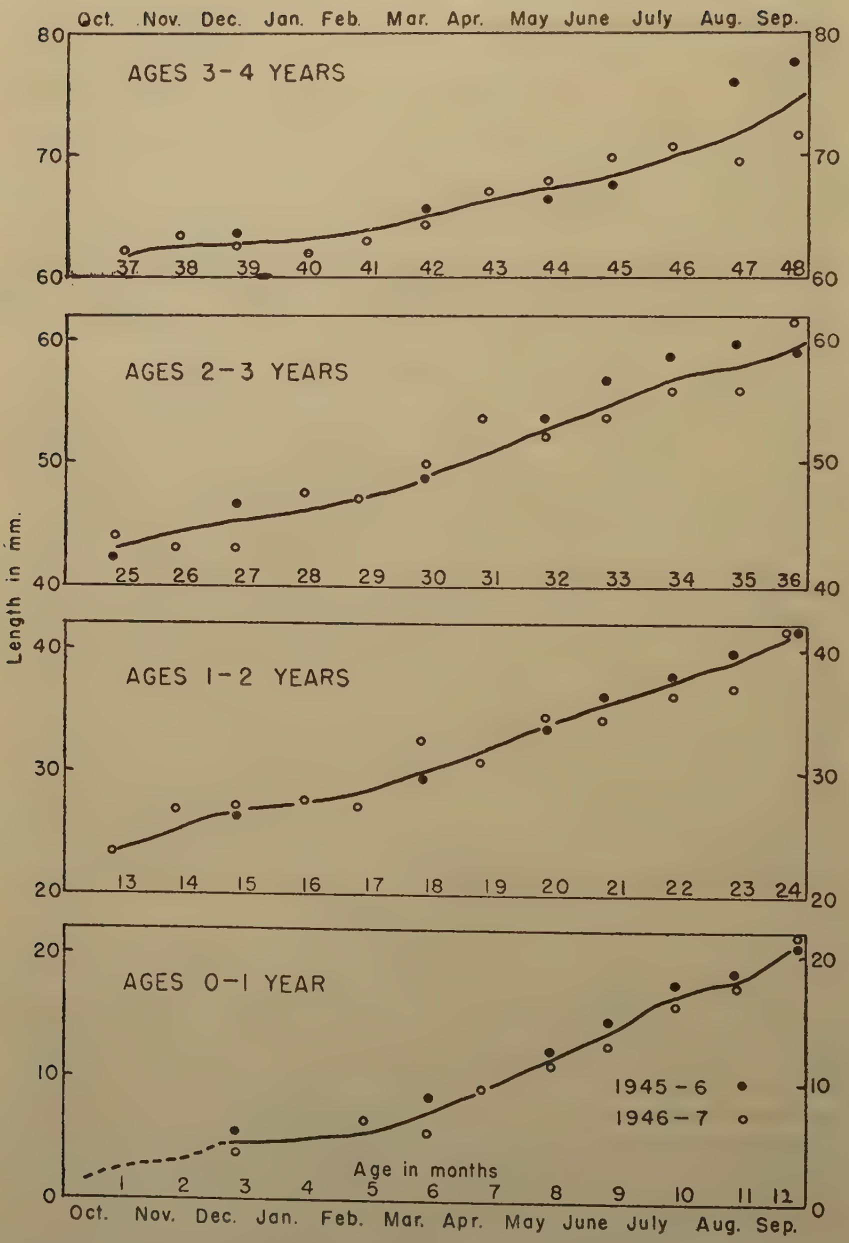


Figure 2. Increase in length of Pismo clam as determined by large samples of the population on the beach at the Scripps Institution of Oceanography taken at monthly intervals during two years. All the clams were returned to the same part of the beach after the measurements had been made. The lower rates of growth in winter and in the month of August are indicated.

total dry organic matter per liter. Some of this is in solution and consequently is unavailable to the mollusks until after it has been converted into particulate or colloidal materials by the action of bacteria or other unicellular organisms. If all were available and if the average adult bivalve filters 22,000 liters of water per year, as estimated by Fox and Coe (1943), there would be a potential supply of 110 g. for each bivalve. Of this the animal stores 1 to 2 g. in the body and discharges as gametes 1 to 3 g. annually (Table II).

TABLE II. MATERIALS INGESTED BY CILIARY-FEEDING BIVALVE MOLLUSKS

	Ingested	*Digested*	
		Living	*Dead (as detritus)*
a. Phytoplankton			
1. Dinoflagellates; 32,000 per l., 12 g.*	+	−	+
2. Diatoms; 32,000 per l., 1.5 g.*	±	±	+
3. Bacteria; 45,000,000 per l., 0.05 g.*	+	+	+
4. Other unicellular and multicellular algae	±	−	+
5. Zoospores and other reproductive cells	+	±	+
b. Zooplankton and nekton			
1. Flagellates	+	±	+
2. Rhizopods and ciliates (including tintinnids)	±	±	+
3. Reproductive products of invertebrates			
Ova and larvae	±	−	+
Spermatozoa	+	+	+
4. Invertebrates and vertebrates	−	−	+
c. Benthic algae and eelgrasses	−	−	+
d. Benthic invertebrates	−	−	+
e. Inorganic and refractory	±	0	0

* Quantity, where given, indicates amount of dry organic matter potentially ingested by one adult individual in one year. Based on computations by Fox and Coe (1943) for a single sea-mussel 70 mm. in length filtering 2.5 l. of water per hour or 22,000 l. per year.

The sign + indicates generally; ± commonly, but only small individuals or species; − rarely or not at all.

All the organic matter in the sea, except those refractory materials such as cellulose, lignin, chitin and others which the mollusks are unable to digest, may be considered as potential nutrients for both mussels and clams. These become available either as the constituents of minute living cells or, more abundantly, as detritus after the death of unicellular and multicellular animals and plants and the disintegration of their cells. Eventually even the refractory organic materials may become available through the action of bacteria.

Therefore, the potential food supply will obviously include: *A*, phytoplankton, consisting of dinoflagellates, diatoms, bacteria and other unicellular and multicellular algae and their reproductive cells; *B*, zooplankton and nekton, including various flagellates, foraminifera and other rhizopods, tintinnids and other ciliates, as well as all invertebrates and vertebrates, with their reproductive cells and larvae;

and for completeness, *C*, all the benthic animals and plants, since they supply no small part of the detritus (Table II). It is obvious that the detritus may be of local origin or brought from a distance by the ocean currents.

PHYTOPLANKTON

The amount of phytoplankton in the water at the Scripps Institution pier is highly variable from year to year and even from month to month. Furthermore, there may be several times as many of the constituent organisms on one day as on the next or preceding day because of the ocean currents which transport the water to and from the area at an average velocity of 4 to 5 miles a day.

In their studies on the nutrition and growth of the sea-mussel, Coe and Fox (1944) have shown that, during the four years 1940 through 1943, those months and those years with the largest dinoflagellate populations were generally conducive to the highest growth rate of the mussels, but that the same correlation did not hold for the diatoms.

Combining the four years gives an average of about 32,000 dinoflagellates per liter, with an approximately equal number of diatoms (Table I, Fig. 1). The dinoflagellates reached a maximum concentration during May, June and July, with a minimum in January and February. In those years there was a sharp decrease in August, the month with the highest temperature of the water, with an increase in September, followed by a gradual decline during the remainder of the year. The diatoms, on the other hand, reached a miximum in August, with a sharp decline toward the end of the year (Fig. 1; Table I).

Over a period of 20 years preceding 1940 there has been a general tendency for both diatoms and dinoflagellates to reach their highest productivity at about the same season of the year. For both groups the period of highest average abundance has been from April to July, with the lowest average in December and January. For several years the lowest average was in August, while in other years the highest average was in that month (Allen, 1940).

Table I, Fig. 1 show that there is some degree of correlation between the monthly averages of both dinoflagellates and diatoms (particularly of dinoflagellates) and the growth rates of mussels and clams, but it should be mentioned again that the dinoflagellates are not utilizable while living and most of the diatoms are too large for intracellular digestion. Such correlations as were found in these experiments are therefore thought to be only indirect and to indicate merely that environmental conditions favorable to the production of large dinoflagellate populations are likewise favorable to the rapid growth of ciliary-feeding mollusks. These studies offer additional evidence of the complexity and variation of the environmental conditions which influence the organisms living on the shores of the oceans.

1. *Dinoflagellates.* The amount of nutrition potentially available (mainly in the form of detritus) from the principal constituents of the phytoplankton has been computed by Fox and Coe (1943) for the sea-mussel, by Coe (1945) for bay-mussel and by Coe (1947) for the Pismo clam. It was estimated that one of these bivalves, on reaching a length of 70 mm., would filter about 60 liters of water during its feeding time each day, or 22,000 liters in a year. With an average dinoflagellate population of 32,000 cells per liter, a total of about 700 million would be available for each individual during the year. If, as estimated, 60 million cells yield 1 g. dry organic matter, then the dinoflagellates presumably ingested in a year would have supplied nearly 12 g. dry organic matter if taken as detritus.

2. *Diatoms.* Similarly, the 32,000 diatoms available per liter would supply only 1.5 g. of organic matter in a year, since these organisms are so small that 500 millions of average size would yield only 1 g. dry organic matter. Most of these, as has been stated, are capable of being digested by the mollusks only after they have become detritus.

3. *Bacteria.* The water in the vicinity of the clam and mussel beds contains an average of about 45 million bacterial cells per liter, and although they are readily ingested and assimilated, they are so minute that they yield but a fraction of a gram of dry organic matter. It is quite possible, however, that they may supply nutritive elements not otherwise available.

4. *Other Unicellular and Multicellular Algae.* Because of their usual cellulose walls, algal cells resist digestion by the bivalve mollusks, but all algae provide nutritive detritus from their cell contents after death and disintegration. Cryptomonads, phytomonads and some of the other flagellates are often abundant and small enough for intracellular digestion. They undoubtedly contribute largely to the bivalves' nutrition.

5. *Zoospores and Other Reproductive Cells of Algae.* These are often very abundant and those that are without cellulose walls are readily ingested and digested. They also become available as detritus.

ZOOPLANKTON AND NEKTON

Only the smallest protozoa are capable of being digested by the intracellular method characteristic of the bivalves, although somewhat larger forms are readily ingested.

1. *Flagellates.* The water often contains vast numbers of various species of flagellates of so minute a size that they are readily available in the living condition. Bruce, Knight and Parke (1940) and others have demonstrated the necessity of these organisms for the normal nutrition of oyster larvae. It seems reasonable to assume that other bivalve mollusks may have the same requirements. Only occasional records are available relative to the numbers of these minute flagellates in the water at different seasons of the year, but it is known that at times they may exceed both the dinoflagellates and diatoms in abundance, with more than a million per liter.

2. *Rhizopods and Ciliates.* With the exception of the flagellates, most of the protozoa are too large for intracellular digestion, but some of them produce reproductive cells which are readily available. Furthermore some of the tintinnids are of the requisite degree of minuteness. All become available after disintegration as detritus.

3. *Reproductive Products of Invertebrates.* Some of the echinoderms and other invertebrates in the vicinity of the clam and mussel beds, including the clams and mussels themselves, set free in the water at certain times of the year vast numbers of spermatozoa. These are admirably adapted for intracellular digestion. The ova and larvae of many species are likewise ingested readily, but since the alimentary canal is deficient in the protease which would be necessary for their digestion, their materials usually become available only after cellular fragmentation. This may occur within the digestive system of the mollusk.

4. *Invertebrates and Vertebrates.* The disintegration of the cells composing the bodies of all the invertebrates and vertebrates of the vicinity, together with their reproductive cells and larvae, provide, after death, an abundant detritus which undoubtedly supplies a large proportion of the bivalves' nutrition. Much of the detritus of local origin is carried elsewhere by the ocean currents, but this is replaced by supplies brought from a distance by the same currents.

BENTHIC ALGAE AND EELGRASSES

Detritus from these sources is always available and doubtless furnishes a considerable proportion of the nutrition after the disintegration of the contents of the cells.

BENTHIC INVERTEBRATES

It seems probable that the vast populations of invertebrates (including the bivalves) along the coast provide a greater quantity of nutritive detritus from their disintegrated bodies, from their gametes and from their feces and excretions than all the other organisms combined.

CHEMICAL AND PHYSICAL CONDITIONS

Chemical Conditions. In localities where there are appreciable changes in the salinity of the sea water, or where toxic substances are added or developed by decomposition or otherwise, the growth rates of the mollusks may be profoundly affected. At the Scripps Institution of Oceanography, however, the rainfall is normally so low as to have little effect on the salinity, and there are no rivers in the vicinity. As a consequence, changes in the chemical composition of the water are too small to influence the growth of the marine invertebrates except in very exceptional seasons. However, there are occasional periods when excessive populations of dinoflagellates, usually in midsummer, may seriously and adversely affect the growth and even the vitality of the mollusks.

Physical Conditions. The graphs in Fig. 1 and the data in Table I indicate considerable differences in the growth rates of both mussels and clams during different months of the year. These appear to be correlated with temperature and solar radiation, as well as with the abundance of phytoplankton. The Scripps Institution is situated on the open coast and is therefore fully exposed to the violence of the surf at the times of occasional storms, usually in the winter.

Whenever these storms occur, not only do the mollusks remain closed and stop feeding, but some of the mussels are torn from their moorings and many of the clams are dislodged by the shifting of the sand, while others are compelled to change their positions in the sand. Since the periods of the storms usually coincide with the season of low temperature, decreased solar radiation and minimum phytoplankton populations, it has seemed impossible to calculate the effects of each of these conditions separately. However, the combined effects of these factors, and presumably others, result in a decreased growth rate of the clam to about one-third and of the mussels to about one-half of the rates characteristic of the most favorable months of the year (Fig. 1, Table I).

TEMPERATURE

Both the mussels and the clams were found to grow most rapidly when the temperature of the water was from 17 to 20° C. Feeding and growth continue, but at a much lower rate, at temperatures of 14° or lower. Experimental evidence was obtained that mussels continue to feed at temperatures as low as 7 to 8° and as high as 25 to 26°, although both these extremes are far beyond the conditions found in nature at La Jolla. In other regions mussels may continue their activities at a temperature of 0° or lower.

Loosanoff (1942) found that on the coast of southern New England the bay-mussel, *Mytilus edulis*, may remain open and active at temperatures ranging from −1.0° to 24.9° C. Feeding continued at or near 0° C. There was a rather close correlation between growth rate and temperature.

Richards (1946) has compared the growth rates of *Mytilus californianus* at La Jolla, California, with those of *M. edulis* at Woods Hole, Massachusetts, with respect to the phytoplankton, temperature and solar radiation. In the latter locality the winter and summer temperatures often differ by as much as 20° C. and at La Jolla by 5° to 8° C. At Woods Hole the maximum rate of growth was in April, followed by a reduced rate in summer and a marked increase in autumn. He concluded that the growth rates in both localities are proportional to the solar radiation but are complicated by changes in food content and temperature of the water.

In more northern localities, both on the coasts of Europe and North America, the annual increment in size is greatly decreased as compared with regions farther south where the temperatures favorable for rapid growth are of longer duration.

On the coast of Norway the growth of the horse-mussel (*Modiola modiolus*) is even slower than that of any of the species mentioned above for localities where the water is considerably warmer. Wiborg (1946) found that the annual increase in length of the young mussels was usually only 6 to 10 mm. a year, with an observed maximum of 15 mm. Individuals 8 years old had an average length of 55 mm., those 10 years old measured 65 to 79 mm., and those 18 years old 98 to 118 mm. Growth was comparatively rapid from July to October but almost ceased as the temperature fell to a minimum during December to April.

SOLAR RADIATION

It is obvious that photosynthesis, and consequently the growth and multiplication of the phytoplankton, is largely dependent on the amount of solar radiation. At the Scripps Institution of Oceanography the incident solar radiation, or insolation, in terms of gram-calories cm^2 per day in the years under consideration reached a maximum in May and July and a minimum in December (Fig. 1; Table I). The abundance of phytoplankton, as well as the growth rates of the bivalve mollusks, follow somewhat similar curves but with many variations for these two years. The similarity of the curves would undoubtedly be increased over a longer term of years.

The graphs (Fig. 1) show that while the solar radiation reached a minimum in December, the lowest populations of phytoplankton occurred one month later, as did also the minimum growth of the sea-mussel. The minimum for the bay-mussel and the Pismo clam was delayed until February, due in part to the influence of storms and the accompanying heavy seas. The maximum solar radiation in May was accompanied by the maximum growth of the sea-mussel. This was followed one month later by the maximum abundance of dinoflagellates and by the most rapid growth of the other two species of bivalves. The maximum population of diatoms in August in these two years likewise followed the maximum solar radiation in July. This was exceptional however, for, as already mentioned, the maximum abundance of diatoms more often occurs in April, following a high period of solar radiation in March and the accumulation of inorganic nutrients during the winter.

SUMMARY

Recent investigations on the nutrition of ciliary-feeding bivalves confirm the evidence that free starch and glycogen are digested in the stomach by enzymes secreted by the cells of the style-sac and incorporated in the style. The gastric juice also contains a small amount of cellulase, but this is inadequate for the complete digestion of dinoflagellates and other phytoplankton with cellulose walls.

All, or nearly all, the rest of the nutrition is derived from the digestion of minute cells and particles within the cells of the digestive organ, supplemented by migratory phagocytes in the alimentary canal. A large proportion of the nutritive materials consists of detritus derived from the disintegration of the cells of unicellular and multicellular plants and animals.

A considerable degree of correlation is found between the growth rate during each month of the year and the food supply, temperature, solar radiation and other environmental conditions.

LITERATURE CITED

Allen, W. E.
1940. Summary of results of twenty years of researches on marine phytoplankton. 6th Pacif. Sci. Congr., *3:* 577–583.

Bruce, J. R., Margery Knight and Mary W. Parke
1940. The rearing of oyster larvae on an algal diet. J. Mar. biol. Ass. U. K., N. S. *24:* 337–374.

CARRIKER, M. R.

1946. Observations on the functioning of the alimentary system of the snail Lymnaea stagnalis appressa Say. Biol. Bull. Woods Hole, *91:* 88–111.

CARRIKER, M. R., AND NELLIE M. BILSTAD

1946. Histology of the alimentary system of the snail *Lymnaea stagnalis appressa* Say. Trans. Amer. micr. Soc., *65:* 250–275.

COE, W. R.

1945. Nutrition and growth of the California bay-mussel (Mytilus edulis diegensis). J. exp. Zool., *99:* 1–14.

1947. Nutrition, growth and sexuality of the Pismo clam (Tivela stultorum). J. exp. Zool., *104:* 1–24.

COE, W. R., AND D. L. FOX

1942. Biology of the California sea-mussel (Mytilus californianus). 1. Influence of temperature, food supply, sex and age on the rate of growth. J. exp. Zool., *90:* 1–30.

1944. Biology of the California sea-mussel (Mytilus californianus). III. Environmental conditions and rate of growth. Biol. Bull. Woods Hole, *87:* 59–72.

FOX, D. L., AND W. R. COE

1943. Biology of the California sea-mussel (Mytilus californianus). II. Nutrition, metabolism, growth and calcium deposition. J. exp. Zool., *93:* 205–249.

FOX, D. L., AND G. W. MARKS

1936. The habitat and food of the California sea mussel; the digestive enzymes. Bull. Scripps Inst. Oceanogr., Tech. Ser., *4* (1): 29–47.

GRAHAM, ALASTAIR

1931. On the morphology, feeding mechanisms, and digestion of *Ensis siliqua* (Schumacher). Trans. roy. Soc. Edinb., *56:* 725–751.

LOOSANOFF, V. L.

1942. Shell movements of the edible mussel, *Mytilus edulis* (L.) in relation to temperature. Ecology, *23:* 231–234.

MANSOUR, K.

1946. Food and the digestive processes of the lamellibranchs. Nature, Lond., *157:* 482.

1946a. The zooxanthellae, morphological peculiarities and food and feeding habits of the Tridacnidae with reference to other lamellibranchs. Proc. Egypt. Acad. Sci., *1:* 1–11.

1946b. Food and digestive organs of lamellibranchs. Nature, Lond., *158:* 378.

MANSOUR, K., AND F. G. ZAKI

1946. The digestive diverticula of Unio prasidens as organs of secretion. Proc. Egypt. Acad. Sci., *2:* 38–44.

MANSOUR-BEK, J. J.

1946. The digestive enzymes of Tridacna elongata Lamk, and Pinctada vulgaris L. Proc. Egypt. Acad. Sci., *1:* 13–20.

1946a. Extracellular proteolytic and lipolytic enzymes of some lamellibranchs. Nature, Lond., *158:* 378–379.

RICHARDS, O. W.

1946. Comparative growth of *Mytilus californianus* at La Jolla, Calif., and *Mytilus edulis* at Woods Hole, Mass. Ecology, *27:* 370–372.

YONGE, C. M.

1926. Structure and physiology of the organs of feeding and digestion in Ostrea edulis. J. Mar. biol. Ass. U. K., N. S. *14:* 295–386.

1932. Notes on feeding and digestion in *Pterocera* and *Vermetus*, with a discussion on the occurrence of the crystalline style in the Gastropoda. Sci. Rep. Gt. Barrier Reef Exped., *1:* 259–281.

1935. On some aspects of digestion in ciliary feeding animals. J. Mar. biol. Ass. U. K., N. S. *20:* 341–346.

1939. The protobranchiate mollusca: a functional interpretation of their structure and evolution. Philos. Trans. roy. Soc. Lond., (B) *230:* 79–147.

1946. Digestion of animals by lamellibranchs. Nature, Lond., *157:* 729.

WIBORG, K. F.

1946. Undersökelser over oskjellet (Modiola modiolus (L)). Fisk Direkt. Skrif., *8:* 1–85.

SULFATE-REDUCING BACTERIA IN MARINE SEDIMENTS[1]

By

CLAUDE E. ZOBELL AND SYDNEY C. RITTENBERG[2]

Scripps Institution of Oceanography
University of California
La Jolla, California

The occurrence of hydrogen sulfide in the waters of certain seas, fjords, and basins, as well as the presence of black or "blue" stinking sediments, can hardly escape the notice of any marine scientist. It is not surprising, therefore, that the eminent oceanographer Murray (Murray and Irvine, 1895) concerned himself with the nature of these materials and postulated their bacterial origin even before bacteriologists had described the specific organisms involved.

It is now accepted that much of the sulfide arises from the bacterial reduction of sulfate. Bacteria which reduce sulfate appear to be very widely distributed in the sea, and it is the purpose of this report to summarize our knowledge of these bacteria with particular reference to their significance in marine sediments, a subject which has been under investigation for several years by the authors and their collaborators at the Scripps Institution.

NOMENCLATURE

Virtually all bacteria and higher plants, as well as a good many animals, may reduce the sulfur in sulfate to the sulfhydryl or disulfide form in the process of synthesizing sulfur-containing proteins, but the term "sulfate reducer" is applied only to those organisms which can utilize sulfates as hydrogen acceptors in their energy-producing reactions. The sulfur in excess of that required for the synthesis of protoplasm is generally reduced to and liberated as hydrogen sulfide by such organisms. As far as is known, only a limited group of bacteria has this unique property. Most of the common bacteria, yeasts, and molds which have been reported to produce traces of hydrogen sulfide from sulfate probably do so by the synthesis of sulfur-containing proteins which, upon the death and autolysis of the cells, are decomposed with the liberation of hydrogen sulfide. The amount of sulfate so reduced to hydrogen sulfide by this process, called "indirect reduction" by Beyjerinck (1900), is extremely small compared to the large amounts reduced to hydrogen sulfide by the true sulfate-reducing bacteria.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 388. This paper is a contribution from A. P. I. Research Project 43A.

[2] Now at the Department of Bacteriology, University of Southern California, Los Angeles, California.

Although there were earlier reports of the microbial reduction of sulfate to hydrogen sulfide, Beyjerinck (1895) is credited with the isolation of the first pure culture of a sulfate reducer originally called *Spirillum desulfuricans*, which was found in ditch water at Delft, Holland. A similar organism, designated *Microspira aestuarii*, which required three per cent salt for its growth, was isolated from mud and water along the North Sea coast by van Delden (1904). A third sulfate-reducing organism called *Vibrio thermodesulfuricans*, allegedly differing from the two named above only in having a higher optimum temperature for growth, was described by Elion (1924).

Baars (1930) regarded all three organisms as varieties of a single species, which he termed *Vibrio desulfuricans*, since he found that certain fresh-water, salt-water, and thermophilic cultures were interconvertible by laboratory acclimatization procedures. These observations were substantiated by Starkey (1938), who also showed that his cultures could be induced to form endospores. He, therefore, proposed the generic name *Sporovibrio* for the entire group of sulfate reducers.

Rittenberg (1941) was unsuccessful in repeated attempts to induce the cultures of sulfate reducers isolated from marine sediments to form spores. Neither could they be adapted to grow in fresh-water media nor at temperatures higher than 40° to 45° C. These observations suggest that at least some of the sulfate reducers found in the sea are sufficiently distinctive to be considered as separate species.

For reasons outlined by the senior author in Bergey's (1948) Manual, it seems best to apply the generic name *Desulfovibrio* to the entire group of sulfate-reducing vibrios, and *Desulfovibrio aestuarii* to the marine forms. Besides *Desulfovibrio aestuarii*, which seems to be quite widely distributed in the sea, there may be several other related species of sulfate reducers in recent or ancient marine sediments and possibly other genera with similar physiology, but information in this regard is too fragmentary for valid conclusions.

MORPHOLOGY

The characteristic form of *Desulfovibrio aestuarii* in young, actively reproducing cultures is that of a slightly twisted spirillum or curved rod ranging in size from 0.6 to 0.8 μ in width by 3 to 4 μ in length.

The young cells may occur in chains of four or more cells attached end to end by plasmodesm, giving such groupings of comma-shaped rods the appearance of long spirilla. Coccus forms and straight rods also occur under certain conditions in young cultures. Older cells may be highly pleomorphic with coccoid forms, rods with swollen ends, and other bizarre forms predominating. Metachromatic granules (probably volutin), sulfur granules, and deposits of iron sulfide have been observed in cells. Beyjerinck (1895), Issatschenko (1929), Baars (1930), and others have noted marked thickenings of older cells or dead cells due to the accumulation of iron sulfide.

While there may be species peculiarities, the morphological variations are attributable primarily to the conditions under which the bacteria have been cultured. The chemical composition, osmotic pressure, pH, and oxygen tension of the medium influence the morphology of sulfate reducers.

In the absence of free oxygen, most young cells are actively motile by means of one or more flagella. Electron microscope photographs of *Desulfovibrio aestuarii* show that these marine species have a single flagellum which generally has a polar insertion. Starkey (1938) found that his sporogenous cultures had peritrichous flagella, the number and location being quite variable.

The *Desulfovibrio* stain fairly well with the common aniline dyes. They are Gram negative.

PHYSIOLOGY

Sulfate reducers are obligate anaerobes. Though not killed rapidly by exposure to oxygen, they fail to reproduce in the presence of free oxygen. Growth is best in a reducing medium, an oxidation-reduction potential of $E_h - 0.1$ to $- 0.3$ volt at pH 7 and 30° C being most favorable for multiplication. In asparagine lactate medium, the first evidence of sulfate reduction observed by Aleshina (1938) was at $E_h + 0.027$ volt. As sulfate reduction proceeded, the oxidation-reduction potential of the medium dropped to $E_h - 0.2$ to $- 0.3$ volt, at which point conditions were optimal. The sulfate reducers studied by Starkey and Wight (1945) grew within the range of pH 5.5 to 8.5 and reduced the E_h from $+ 0.3$ volt to $- 0.2$ or $- 0.3$ volt. The pH range of the marine species is 5.4 to 9.0, the optimum being pH 6 to 8.

Under favorable conditions certain cultures may produce as much as 1,500 mgm/L. of hydrogen sulfide. According to Baars (1930), halophilic strains produce considerably more hydrogen sulfide than fresh-water strains—his cultures of the former producing a maximum of 1,448 mgm/L. as compared with only 794 mgm/L. for the latter. Rittenberg (1941) found that the rate and the total amount of hydrogen sulfide produced by *Desulfovibrio aestuarii* varied with the concentration of salt in the medium. In general, increasing the salt content within the range of 2 to 4% diminished the rate but increased the total amount of hydrogen sulfide produced. This may be illustrated by the following results:

Period of incubation	*Hydrogen sulfide formed in media containing*		
	2% NaCl	*3% NaCl*	*4% NaCl*
119 hours	56 mgm/L.	37 mgm/L.	8 mgm/L.
143 hours	370 mgm/L.	234 mgm/L.	30 mgm/L.
167 hours	520 mgm/L.	513 mgm/L.	126 mgm/L.
191 hours	622 mgm/L.	660 mgm/L.	518 mgm/L.
239 hours	692 mgm/L.	809 mgm/L.	902 mgm/L.
287 hours	715 mgm/L.	826 mgm/L.	1013 mgm/L.

Sulfur, sulfite, thiosulfate, and tetrathionate are also reduced to hydrogen sulfide by sulfate reducing bacteria. These compounds are assumed to be intermediates in the reduction of sulfate and are assumed to be reduced more readily than sulfate, since they do not accumulate in the medium.

It is a moot question whether sulfate is ever reduced only to and deposited as sulfur in an anaerobic environment, although recently Datta (1943) described an organism as a new species which was claimed to have this ability. Sulfur is often found in cultures or cells of sulfate-reducing bacteria, but there is always some question as to whether it is formed directly from the reduction of sulfate or from the oxidation of sulfide upon coming into contact with oxygen or other oxidizing agents. It is possible, as suggested by Zil'berman (1940), that sulfite, which is believed to be an intermediate product in the reduction of sulfate, reacts with hydrogen sulfide to produce free sulfur: $H_2SO_3 + 2H_2S \rightarrow 3S + 3H_2O$

Apart from sulfate and the intermediate inorganic sulfur compounds already mentioned, no other hydrogen acceptors have been demonstrated for this group of bacteria. Several organic sulfur compounds, phosphate, selenate, tellurate, and peptone were tested in this connection (Rittenberg, 1941) with negative results. Nitrate, although assimilated as a source of nitrogen, was not utilized as a hydrogen acceptor in the energy-yielding reactions of the sulfate reducers.

There are strains of sulfate-reducing bacteria which can utilize nearly any type of organic matter as hydrogen donors. Energy can apparently be derived by sulfate reduction and the concomitant oxidation of cellulose (Rubentschik, 1928), chitin and glucosamine (Aleshina, 1938), beef and mutton tallow (Seliber, 1928, 1945), aliphatic and aromatic hydrocarbons (Rosenfeld, 1947), and other resistant materials.

Simple compounds like lactate and asparagine are the best sources of carbon or energy for most cultures of sulfate-reducing bacteria. A large variety of mono- and disaccharides, alcohols, aldehydes, aliphatic acids, and amino acids are metabolized by some species.

The types of materials used vary somewhat from strain to strain. Baars (1930) described *Desulfovibrio rubentschikii* as a new species of sulfate reducer primarily upon a basis of its ability to attack fatty acids and sugars not used by other species.

HYDROGEN UTILIZATION

A large body of data indicate that at least certain sulfate reducers can obtain energy by the oxidation of molecular hydrogen and even live autotrophically by this mechanism. The work of Nikitinsky (1907), Kroulik (1913), and Niklewski (1914) gave presumptive evidence of hydrogen uptake by sulfate reducers, but it remained for Stephenson and Stickland (1931) to isolate, from river mud, a pure culture which reduced sulfate to hydrogen sulfide with the consumption of molecular hydrogen:

$$H_2SO_4 + 4H_2 \rightarrow H_2S + 4H_2O + 56{,}750 \text{ calories.}$$

This culture, designated Strain 182, also catalyzed the reduction of sulfite and thiosulfate with molecular hydrogen. The demonstration of hydrogen uptake was obtained with resting cells using the Warburg technique, and there was no proof of carbon dioxide uptake or cell synthesis under these conditions. Wight and Starkey (1945), however, demonstrated the ability of sulfate reducers to live autotrophically by obtaining their energy requirements from the oxidation of molecular hydrogen and their carbon from bicarbonate or carbon dioxide. Certain other cultures studied by Starkey and Wight (1945) utilized hydrogen in lactate medium but not in a purely mineral medium.

Several different pure or partially purified cultures of sulfate-reducing bacteria, isolated from marine mud, were reported by ZoBell (1947a) to consume hydrogen while producing hydrogen sulfide from sulfate in a strictly mineral medium. The ratio of the amount of hydrogen consumed to the amount of hydrogen sulfide produced exceeded the theoretical value of 4.0 which would result if hydrogen were oxidized quantitatively to hydrogen sulfide and water. This suggests that part of the hydrogen may have been used to reduce carbon dioxide in the synthesis of cell substance.

OTHER CULTURAL REQUIREMENTS

Considerable importance has been attached to the salinity requirements of sulfate-reducing bacteria, because this has been employed as the main criterion for distinguishing the marine or halophilic species, *Desulfovibrio aestuarii*, from other species. Baars (1930) was able to acclimatize a fresh-water strain to grow in a medium containing 3% salt, and therefore he concluded that the strains were identical. It should be pointed out that the fresh-water culture Baars employed produced some hydrogen sulfide in his 3% salt medium before the adaptation procedure. In this instance at least "adaptation" may have been a selection process. Rittenberg (1941), however, could not induce halophilic strains which he isolated from marine sediments to grow in nutrient fresh-water medium. They grew normally, though with decreased sulfide production, when the sea water medium, in which they were most active, was diluted 1 : 1 with fresh-water medium. A few cultures were induced to grow in a nutrient medium containing only 25 parts of sea water to 75 parts of fresh water, but only a little hydrogen sulfide was produced. The lack of adaptability was also reflected in the abnormal morphology of the sulfate reducers which were induced to grow in media containing less than 50% sea water.

The salinity picture is further complicated by the observations of other investigators. Strains apparently indistinguishable from *Desulfovibrio desulfuricans* and *Desulfovibrio aestuarii* in all other respects have been described with a salt optimum of 8% and a growth range between traces and 20% sodium chloride (Chait, 1924); with an optimum of 5% (Rubentschik, 1928); with salt ranges of traces to 1%, traces to 5%, and 1% to 3% (Gahl and Anderson, 1928); with a range of 1% to 18% (Ginsburg-Karagitscheva, 1933); and so on. It is entirely possible that there are obligate halophiles, obligate fresh-water strains, and a host of intermediate forms with varying ranges of salt tolerances. Any attempt to classify these various forms into separate species, however, should logically wait until some knowledge is obtained concerning the mechanism of salt adaptation and the factors controlling the salt ranges of growth.

Similarly, various temperature ranges and optima have been reported for different cultures of sulfate-reducing bacteria. His being able to develop cultures, from the same material, having an optimum temperature of either 30° or 55° C., led Baars (1930) to regard Beyjerinck's (1895) *desulfuricans* (optimum temperature of 30°) and Elion's (1924) *thermodesulfuricans* (optimum temperature of 55°) as adaptation forms of the same plurivalent or thermally adaptable type. However, since Baars did not start with single cell isolations, he may have been producing enrichment cultures of one type or the other of organisms of different potentialities initially present by the selective temperature of incubation. Rittenberg (1941) was unable to acclimatize pure cultures of *Desulfovibrio aestuarii* isolated from marine sediments to reduce sulfate at temperatures higher than 40° to 45° C.

OCCURRENCE IN THE SEA

Sulfate-reducing bacteria appear to be abundantly and widely distributed in the marine environment, especially in bottom deposits. Such bacteria are occasionally found in samples of sea water, where their presence may be of doubtful significance, since sulfate reducers are active only in strictly anaerobic environments. In the wake of a prolific diatom bloom or other circumstances resulting in the occurrence of enough decaying organic matter to consume all of the dissolved oxygen, or in stagnant water basins, conditions in sea water may be suitable for the activity of sulfate reducers. Such conditions prevail in the Black Sea, the water of which contains no free oxygen at depths exceeding 200 meters (Sverdrup, *et al.*, 1942). Water below the euphotic zone in the Caspian Sea, Norwegian fjords, and certain localized areas in the ocean are also sufficiently anaerobic at times to favor the growth of sulfate reducers. The activities of sulfate reducers are largely responsible for maintaining anaerobic conditions by the production of hydrogen sulfide which depletes free oxygen.

According to Daniltchenko and Tschigirine (1926), the concentration of hydrogen sulfide near the bottom of the Black Sea, where sulfate reducers are most active, is sometimes as high as 3,000 mgm/L. Copenhagen (1934) described an area approximately 25 by 200 miles in the Atlantic Ocean off Walvis Bay, South Africa, where hydrogen sulfide is produced periodically by sulfate reducers in quantities sufficient to be lethal to flora and fauna almost to the surface of the overlying water.

It was in the Black Sea where sulfate-reducing bacteria were first demonstrated by Zelinski (1893). Issatschenko (1924, 1929) and Ravich-Sherbo (1930) found vigorous sulfate reducers in virtually all Black Sea water samples examined. In the Caspian Sea, Butkevich (1938) found up to two million bacteria per ml. of water, many of which were sulfate reducers. Sulfate reducers were found in the Salton Sea by Brannon (1914), in the Mediterranean Sea by Bertel (1935), and in Russian limans by Rubentschik (1928). The freshening of certain limans in the Odessa region during the rainy season was found by Rubentschik and Goicherman (1936) to be followed by a ten- to thousand-fold increase in the abundance of sulfate reducers.

In bottom deposits, where bacteria are generally most abundant, ZoBell (1938) demonstrated from 1,000 to 10,000 sulfate reducers per gram. Large numbers of sulfate reducers were found by Bavendamm (1932) in nearly all samples of calcareous mud in the region of the Bahama Islands. A good many of the bacteria recovered from deep sea bottom deposits by Ginsburg-Karagitschewa and Rodionowa (1935) were sulfate reducers.

Rittenberg (1941) found viable sulfate reducers in about half of 450 2- to 3-gram samples (wet weight) from 160 different stations off the coast of California and Lower California. Profile series showed that the abundance decreased with the depth of the sediment, sulfate reducers being detected in only 10% of the samples from depths exceeding 60 inches, as compared with 32% positive at depths of 12 to 60 inches, and 64% positive in the topmost 12 inches of sediment. Sulfate reducers were detected in 92% of the samples of silty sand, 86% of the sandy silt samples, 65% of the silty mud samples, and in 51% of the samples of clayey mud.

The presence of these bacteria, together with the occurrence of sulfide and the decreasing concentration of sulfate with depth in the interstitial water in certain sediments, suggests that the sulfate reducers have been active *in situ.*

OCCURRENCE IN ANCIENT MARINE SEDIMENTS

Bastin (1926) demonstrated the presence of sulfate-reducing bacteria in 28 oil-field waters, believed to have a marine origin, out of 30 sampled in Illinois and in 15 out of 37 sampled in California. He attributed the generally low content of sulfate in oil-well waters to the activity of the bacteria. The samples of oil-well waters were obtained from ancient marine sediments at depths ranging from 400 to 1,866 feet. Similar observations were made by Gahl and Anderson (1928), who found sulfate reducers in 17 out of 40 oil-well waters sampled in California fields from depths as great as 3,090 feet. The isolated organisms grew at temperatures ranging from 28° to 50° C. The temperature of the oil-well waters was 44° to 47° C.

Sulfate reducers were found in the water from six producing oil wells in Illinois fields by Bastin and Greer (1930). The waters were rich in hydrogen sulfide. In oil-well waters containing no sulfide, sulfate reducers were absent. Further evidence that sulfate-reducing bacteria are present and active in waters associated with petroleum deposits has been presented by Ginter (1930, 1934). He points out that carbonate and sulfide in many brines increased as sulfate decreased, presumably due to the reduction of sulfate to hydrogen sulfide and the simultaneous oxidation of carbon compounds to carbon dioxide which appears as carbonate or bicarbonate.

Besides demonstrating the occurrence of sulfate reducers in water from several oil wells, Ginsburg-Karagitscheva (1933) reported the presence of such bacteria in 14 out of 15 samples of oil sands from Grozny oil-field formation. The bacteria were believed to be indigenous species responsible for the abundance of hydrogen sulfide, the low concentration or lack of sulfate, and the modification of the associated crude oil. Baier (1937) reasoned that sulfate reducers could be active at the oil-water interface where petroleum constituents provide the carbon and energy necessary for the growth of the bacteria.

While the possibilities of surface contamination have not been ruled out, it is significant that sulfate reducers are generally the predominant bacteria detected in oil-well waters. They grow best or exclusively at the temperature and salinity of the water in which they were found, and it has been established that the sulfate reducers can obtain all of their food requirements from the reservoir fluids in which they are found. This circumstantial evidence, plus the fact that appreciable numbers of sulfate reducers have been recovered from oil wells long after anything has been introduced into the wells, strongly suggests that if the bacteria are not actually species indigenous to the formation, they have been multiplying in the oil wells.

ZoBell (1946) detected living sulfate-reducing bacteria in aseptically collected samples of sulfur limestone anhydrite from a depth of 1,560 feet. Cultures which reduced sulfate at temperatures ranging from 75° to 85° C have been observed. Recent observations indicate that the activity of sulfate reducing bacteria is promoted by high hydrostatic pressures such as prevail in the ocean or in petroleum deposits at depths of several thousand feet.

IMPORTANCE OF SULFATE REDUCERS IN SEDIMENTS

There are several ways in which the activities of sulfate-reducing bacteria may modify conditions in marine sediments both recent and ancient. Besides contributing to the sulfur cycle in sediments (Galliher, 1933) by converting sulfate into sulfide and possibly to free sulfur, they transform various kinds of organic matter, consume molecular hydrogen, affect the pH and E_h of their environment, and otherwise influence chemical or physico-chemical conditions.

The free sulfur quite commonly found by Buchanan (1891) in marine bottom deposits was believed to result from sulfides coming in contact with oxygen in the overlying water. He was mistaken, though, in believing that sulfate was reduced to sulfide by the digestive secretions of animals. Hunt (1915) has attributed the origin of sulfur deposits in Sicily to the bacterial reduction of sulfate, and the Louisiana and Texas Gulf Coast sulfur deposits were believed by Ridgway (1930) to result from the bacterial reduction of gypsum. Ahlfeld (1937), Murzaiev (1937), and Sturm (1937) agree that certain sulfur deposits in Russia resulted from the oxidation of hydrogen sulfide that could have come only from the bacterial reduction of sulfate. Associated with sulfur deposits on the east coast of India, Iya and Sreenivasaya (1945 a, b) found sulfate-reducing bacteria which were thought to form sulfur.

Sulfate reducers may quite definitely contribute to the formation of metallic sulfide deposits. Murray and Irvine (1895) attributed the formation of ferrous sulfide deposits in black marine mud to the activity of sulfate-reducing bacteria, which were also believed to account for the loss of sulfate and an increase in the carbonate content, the latter resulting from the oxidation of organic matter by bacteria. The increased carbonate content in the mud examined on the CHALLENGER Expedition was equivalent to the decreased sulfate content.

Issatschenko (1929) demonstrated the formation of iron sulfide inside the cells of sulfate-reducing bacteria, which were credited with being responsible for the deposition of pyrite in the Black Sea and the Sea of Azov. Similar deposits of iron sulfide, attributable to the activities of sulfate reducers, have been reported by Doss (1912) in the Mediterranean, Caspian, and Black Sea, by Ellis (1925) in the Clyde Sea Aestuary, by Kindle (1926) in the Bay of Fundy, by Strøm (1939) in Norwegian fjords, and by Copenhagen (1934) in Walvis Bay off the South African Coast. Manganese sulfide may be precipitated in a like manner. The formation of copper sulfide in marine sediments has been attributed to the activities of sulfate-reducing bacteria by Schneiderhöhn (1923), Trask (1925), Thiel (1926), and Bastin (1933). Thiel (1927) explained the formation of bauxite and other aluminum oxides from aluminous silicates as being partly due to the activities of sulfate-reducing bacteria.

Not only is sulfate in interstitial water reduced, but sulfate rocks in consolidated sediments may be attacked by sulfate-reducing bacteria. Gypsum, anhydrite, or other sulfate rocks may be slowly dissolved under conditions which are favorable for the activity of sulfate-reducing bacteria (ZoBell, 1947b). The dissolution of sulfate rocks is favored by low pH. The conversion of the strongly acidic sulfate radicle to neutral sulfur or to the only slightly acidic sulfide radicle tends to increase the pH, but the oxidation of organic compounds required as a source of energy for the sulfate-reducing bacteria generally yields acidic products, so the over-all effect may be to lower the pH. In one series of experiments, for example, Rittenberg (1941) found the final reaction of unbuffered media to be pH 5.8 to 5.9 as a result of the activity of sulfate reducers for 35 days in media having initial reactions of pH 6.0, 7.0, and 7.8, respectively.

The commonly reported occurrence of sulfate-reducing bacteria in the reservoir fluids in petroliferous marine sediments suggests that these bacteria may play a role in the formation or modification of petroleum. That such bacteria may modify the chemical composition of certain oil-field waters is quite generally accepted. The ability of some of them to attack petroleum hydrocarbons has been established (Tausson and Aleshina, 1932; Tausson and Vesselov, 1934; Novelli and ZoBell, 1944; Rosenfeld, 1947). The formation of liquid and waxy hydrocarbons by sulfate reducers has been reported by Jankowski and ZoBell (1944). Fatty acids ranging from acetic to stearic were also assimilated with the formation of small amounts of hydrocarbons in the cell substance of sulfate reducers.

Sulfate-reducing bacteria may have contributed greatly to the formation and migration of petroleum in ancient marine sediments by releasing oil from oil-bearing materials and promoting its flow. ZoBell (1947b) has shown that this may be done by dissolving either carbonate or sulfate rocks on which oil is adsorbed, by producing carbon dioxide which decreases the viscosity of oil, by the bacteria's displacing oil films physically, or by their producing detergents or surface active substances.

These bacteria are also of significance in the anaerobic corrosion of iron, a field actively investigated by Wight and Starkey (1945) and others.

It may also be mentioned that these organisms have been isolated from mineral waters of various spas and resorts in Europe and America (Gosling, 1904; Rank, 1907). One might speculate as to whether the sulfate reducers contribute to the "curative" properties of these waters or merely to their odors.

SUMMARY

Desulfovibrio aestuarii is described as the type species of marine bacteria which reduce sulfates as hydrogen acceptors in their energy-yielding reactions, although there may be other genera and species. Such bacteria are physiologically active only in the absence of free oxygen and they tend to maintain or create reducing conditions by producing hydrogen sulfide, particularly in stagnant water basins and in sedimentary deposits. Sulfate-reducing bacteria oxidize a large variety of organic compounds and some can utilize molecular hydrogen as an energy source.

Sulfate-reducers occur quite widely and abundantly in marine bottom deposits. They have also been found at appreciable depths in ancient marine sediments where they may play a part in petroleum formation. Besides affecting the pH, E_h, and organic content of bottom deposits, sulfate-reducers may be responsible for the formation of various metallic sulfides and the modification of other minerals in marine sediments.

REFERENCES

AHLFELD, FRIEDRICH

1937. Genesis of sulphur deposits of the U. S. S. R. Econ. Geol., *32:* 974–976.

ALESHINA, W. I.

1938. Decomposition of chitin by sulphate-reducing bacteria and changes in oxidation-reduction conditions during the process of reduction of sulphates. Mikrobiologiia, *7:* 850–859.

BAIER, C. R.

1937. Bakteriologische Erdölstudien. Kieler Meeresforsch., *2:* 149–156.

BAARS, J. K.

1930. Over Sulfaatreductie door Bacteriën. Dissertation. Delft, Holland. 164 pp.

BASTIN, E. S.

1926. The problem of the natural reduction of sulphates. Bull. Amer. Ass. Petrol. Geol., *10* (2): 1270–1299.

1933. The chalcocite and native copper types of ore deposits. Econ. Geol., *28:* 407–446.

BASTIN, E. S. AND F. E. GREER

1930. Additional data on sulphate-reducing bacteria in soils and waters of Illinois oil fields. Bull. Amer. Ass. Petrol. Geol., *14* (1): 153–159.

BAVENDAMM, W.

1932. Die mikrobiologische Kalkfällung in der tropischen See. Arch. Mikrobiol., *3:* 205–276.

BEYJERINCK, M. W.

1895. Ueber *Spirillum desulfuricans* als Ursache von Sulfatreduction. Zbl. Bakt., II Abt., *1:* 1–9, 49–59, 104–114.

1900. Schwefelwasserstoffbildung in den Stadtgräben und Aufstellung der Gattung Aërobacter. Zbl. Bakt., II Abt., *6:* 193–206.

BERGEY, D. H., *et al.*

1948. Manual of Determinative Bacteriology. Williams & Wilkins Co., Baltimore. 6th Ed. (pp. 207–209) 1529 pp.

BERTEL, R.

1935. Les bactéries marines et leur influence sur la circulation de la matière dans la mer. Bull. Inst. océanogr. Monaco, No. 672: 1–12.

BRANNON, M. A.

1914. The action of Salton Sea water on vegetable tissues. Publ. Carneg. Inst., No. 193: 71–78.

Buchanan, J. Y.
1891. On the occurrence of sulfur in marine muds and nodules, and its bearing on their mode of formation. Proc. roy. Soc. Edinb., *18:* 17–39.

Bunker, H. J.
1936. A review of the physiology and biochemistry of the sulphur bacteria. His Majesty's Stationery Office, London. 48 pp.

Butkevich, V. S.
1938. On the bacterial population of the Caspian and Azov seas. Mikrobiologiia, *7:* 1005–1021.

Chait, S. S.
1924. Ber. wiss. Forsch. Inst. Odessa, *1:* no. 10–11. (Cited by Baars, 1930).

Copenhagen, W. J.
1934. Occurrence of sulphides in certain areas of the sea bottom on the South African Coast. Fish. Mar. biol. Surv., Un. S. Afr., Rep. No. 11: 3–18.

Daniltchenko, P. and N. Tschigirine
1926. Sur l'origine de l'hydrosulfure dans la mer Noire. Trav. Lab. zool. Sébastopol, Ser. 2, No. 10: 141–191.

Datta, S. C.
1943. Production of elementary sulphur by reduction of sulphate through bacterial agency. Curr. Sci., *12:* 305.

Doss, Bruno
1912. Melnikowit, ein neues Eisenbisulfid, und seine Bedeutung für die Genesis der Kieslagerstätten. Z. prakt. Geol., *20:* 453–483.

Elion, L.
1924. A thermophilic sulphate-reducing bacterium. Zbl. Bakt., II Abt., *63:* 58–67.

Ellis, D.
1925. J. Roy. Tech. Coll., Glasgow. p. 144. (Cited by Bunker, 1936).

Gahl, Rudolf and Belle Anderson
1928. Sulphate reducing bacteria in California oil waters. Zbl. Bakt., II Abt., *73:* 331–338.

Galliher, E. W.
1933. The sulfur cycle in sediments. J. sediment. Petrol., *3:* 51–63.

Ginsburg-Karagitscheva, T. L.
1933. Microflora of oil waters and oil-bearing formations and biochemical processes caused by it. Bull. Amer. Ass. Petrol. Geol., *17:* 52–65.

Ginsburg-Karagitscheva, T. L. and K. Rodionowa
1935. Beitrag zur Kenntnis der im Tiefseeschlamm stattfindenden biochemischen Prozesse. Biochem. Z., *275:* 396–404.

Ginter, R. L.
1930. Causative agents of sulphate reduction in oil-well waters. Bull. Amer. Ass. Petrol. Geol., *14:* 139–152.
1934. Sulphate reduction in deep subsurface water. In: Problems of Petroleum Geology. Amer. Ass. Petrol. Geol., Tulsa (pp. 907–925). 1073 pp.

Goslings, N.
1904. Ueber schwefelwasserstoffbildende Mikroben in Mineralwässern. Zbl. Bakt., Abt. II, *13:* 385–394.

HUNT, W. F.
1915. The origin of the sulphur deposits of Sicily. Econ. Geol., *10:* 543–579.
ISSATSCHENKO, B. L.
1924. Sur la fermentation sulfhydrique dans la mer Noire. C. R. Acad. Sci., Paris, *178:* 2204–2206.
1929. Zur Frage der biogenischen Bildung des Pyrits. Int. Rev. Hydrobiol., *22:* 99–101.
IYA, K. K. AND M. SREENIVASAYA
1945a. Studies in the formation of sulphur at Kona (Masulipatam). Part I. Curr. Sci., *14:* 243–244.
1945b. Studies in the sulphur formation at Kona, Masulipatam. Part II. Curr. Sci., *14:* 267–269.
JANKOWSKI, G. J. AND C. E. ZOBELL
1944. Hydrocarbon production by sulfate-reducing bacteria. J. Bact., *47:* 447.
KINDLE, E. M.
1926. Notes on the tidal phenomena of Bay of Fundy Rivers. J. Geol., *34:* 642–652.
KROULIK, ALOIS
1913. Über thermophile Zellulosevergärer. Vorläufige Mitteilung. Zbl. Bakt., II Abt., *36:* 339–346.
MURRAY, J. AND R. IRVINE
1895. On the chemical changes which take place in the composition of the sea water associated with blue muds on the floor of the ocean. Trans. roy. Soc. Edinb., *37:* 481–507.
MURZAIEV, P. M.
1937. Genesis of some sulphur deposits of the U. S. S. R. Econ. Geol., *32:* 69–103.
NIKITINSKY, J.
1907. Die anaërobe Bindung des Wasserstoffes durch Mikroorganismen. Zbl. Bakt., II Abt., *19:* 495–499.
NIKLEWSKI, BRONISLAW
1914. Über die Wasserstoffaktivierung durch Bakterien unter besonderer Berücksichtigung der neuen Gattung *Hydrogenomonas agilis*. IV. Zbl. Bakt., II Abt., *40:* 430–433.
NOVELLI, G. D. AND C. E. ZOBELL
1944. Assimilation of petroleum hydrocarbons by sulfate-reducing bacteria. J. Bact., *47:* 447–448.
RANK, A.
1907. Dissertation, Zurich. (Cited by Baars, 1930.)
RAVICH-SHERBO, J.
1930. To the question of bacterial thin layer in the Black Sea according to the hypothesis of Prof. Egounoff. Trav. Sta. biol. Sébastopol, *2:* 127–141.
RIDGWAY, R. H.
1930. Sulphur. General Information. U. S. Bur. Mines, I. C. *6329:* 1–55.
RITTENBERG, S. C.
1941. Studies on marine sulphate-reducing bacteria. Dissertation No. 650, Univ. Calif., La Jolla. 115 pp.

ROSENFELD, W. D.

1947. Anaerobic oxidation of hydrocarbons by sulfate-reducing bacteria. J. Bact., *54:* 664–665.

RUBENTSCHIK, L. I.

1928. Über Sulfatreduktion durch Bakterien bei Zellulosegärungsprodukten als Energiequelle. Zbl. Bakt., II Abt., *73:* 483–496.

RUBENTSCHIK, L. I. AND D. G. GOICHERMAN

1936. The influence of a decrease in salt-content in limans on the microflora of medicinal muds. Arkh. biol. nauk, *43* (ser. 2–3): 217–227.

SCHNEIDERHÖHN, H.

1923. Chalkographische Untersuchung des Mansfelder Kupferschiefers. N. Jb. Min. Geol. Paläont., Beil., *47:* 1–38.

SELIBER, G. L.

1928. La réduction des sulphates par des microorganismes en présence des graisses. C. R. Soc. biol., *99:* 544–546.

1945. Concerning the production of hydrocarbons from fats by microorganisms. Mikrobiologiia, *14:* 254–259.

STARKEY, R. L.

1938. A study of spore formation and other morphological characteristics of *Vibrio desulfuricans.* Arch. Mikrobiol., *9:* 268–304.

STARKEY, R. L. AND K. M. WIGHT

1945. Anaerobic corrosion of iron in soil. Tech. Rep. Distrib. Comm., No. 1945, Amer. Gas Ass. 108 pp.

STEPHENSON, MARJORY AND L. H. STICKLAND

1931. Hydrogenase. II. The reduction of sulphate to sulphide by molecular hydrogen. Biochem. J., *25:* 215–220.

STRØM, K. M.

1939. Land-locked waters and the deposition of black muds. In: Recent Marine Sediments. Amer. Ass. Petrol. Geol., Tulsa. (pp. 356–372) 736 pp.

STURM, L. D.

1937. A study of the microflora of sulfur-containing rocks. Microbiologiia, *6:* 481–497.

SVERDRUP, H. U., M. W. JOHNSON, AND R. H. FLEMING

1942. The Oceans. Prentice-Hall, Inc., New York. 1087 pp.

TAUSSON, W. O. AND W. A. ALESHINA.

1932. Über die bakterielle Sulfatreduktion bei Anwesenheit der Kohlenwasserstoffe. Mikrobiologiia, *1:* 229–261.

TAUSSON, W. O. AND I. J. VESSELOV

1934. On the bacteriology of the decomposition of cyclical compounds at the reduction of sulphates. Mikrobiologiia, *3:* 360–369.

THIEL, G. A.

1926. The Mansfeld Kupferschiefer. Econ. Geol., *21:* 299–300.

1927. The enrichment of bauxite deposits through the activity of microörganisms. Econ. Geol., *22:* 480–493.

TRASK, P. D.

1925. The origin of the ore of the Mansfeld Kupferschiefer, Germany. A review of the current literature. Econ. Geol., *20:* 746–761.

van Delden, A.

1904. Beitrag zur Kenntnis der Sulfatreduktion durch Bakterien. Zbl. Bakt., II Abt., *11:* 81–94, 113–119.

Wight, K. M. and R. L. Starkey

1945. Utilization of hydrogen by sulfate-reducing bacteria and its significance in anaerobic corrosion. J. Bact., *50:* 238.

Zelinski, N. D.

1893. On hydrogen sulfide fermentation in the Black Sea and the Odessa estuaries. Proc. Russ. Phys. Chem. Soc., *25*: 298–303.

Zil'berman, Ya. I.

1940. Reaction of hydrogen sulfide with salts of sulfurous acid. J. gen. Chem., Moscow, *10:* 1257–1265.

ZoBell, C. E.

1938. Studies on the bacterial flora of marine bottom sediments. J. sediment. Petrol., *8:* 10–18.

1946. Functions of bacteria in the formation and accumulation of petroleum. Oil Wkly., *120* (12): 30–36.

1947a. Microbial transformation of molecular hydrogen in marine sediments, with particular reference to petroleum. Bull. Amer. Ass. Petrol. Geol., *31:* 1709–1751.

1947b. Bacterial release of oil from oil-bearing materials. World Oil, *126* (13): 36–47; *127* (1): 35–41.

RADIOACTIVITY OF OCEAN SEDIMENTS
VII. RATE OF DEPOSITION OF DEEP-SEA SEDIMENTS

By

WM. D. URRY

Geophysical Laboratory
Carnegie Institution of Washington

INTRODUCTION

Many of the problems of past geological history and of present processes of sedimentation require a knowledge of the rate of accumulation of marine sediments. A very important step in the determination of the rate of deposition of a sediment is the measurement of a time interval. This is equally true whether use be made of the *stratigraphic* method with core samples or the *supply* method based upon the rate of supply of sedimentary material to the sea. The stratigraphic method gives rates for individual localities, whereas the supply method yields averages for large areas over long periods of geological history.

The methods previously employed to estimate the rates of deposition of ocean sediments are well summarized in the chapter on "Marine Sedimentation" in *The Oceans* (13). It is only necessary here to emphasize the subjectivity of these methods and their limitations. Schott (12) arrived at the mean rates of deposition of the three main types of ocean sediments by assuming that the pelagic foraminifera, *G. menardii*, began to accumulate in the deposits of the equatorial Atlantic 20,000 years ago. *G. menardii* were absent or nearly absent while the lower strata represented in the short cores collected by the Meteor were being deposited. The choice of such a time interval has no basis in fact, even though it may be guided by the estimates of various authors of the beginning of postglacial time over the land areas of the northern hemisphere. Further, Schott's estimates are likely to be in error because no correction was made for distortion of the sediment in the coring of the sample. This distortion was probably quite appreciable in view of the narrow liner employed and the fact that the boundary of particular interest was often close to the bottom of the core. But quite apart from the preliminary nature of such estimates, generally well recognized by investigators making use of the data, no analysis of the changes in the rate of deposition in the past is possible from such measurements, and application of such averages to dating specific vertical sections of the ocean bottom is certainly not advisable.

Average rates, based on supply methods, are dependent, among other things, on the accuracy with which the extent of erosion over a long period of geological time can be estimated and on the estimate of the area over which the chemical and detrital deposition occurs. For some purposes such averages obtained by Kuenen (4), Revelle and Shepard (11), and Lohmann (5) are useful, but these figures can never be more than averages for large areas and for long periods of time.

A series of measurements of the radium content of core samples of marine sediments by Piggot and Urry (8, 9, 10), Urry and Piggot (17) and Urry (16) has demonstrated that properties of the radioelements can be employed to date the events that have left their records in the ocean bottom. The radioactive method of dating possesses several advantages. It is free from any consideration of the water content or distortion in the core sample, or of the particular rate of deposition at the site of the sample or the temporal variations of this rate. Piggot and Urry (10) have discussed the application of this method to core samples from the North Atlantic, the Caribbean Sea, and the North Pacific.

It should be a simple matter to calculate the rate of deposition, and the changes in this rate with time, from measurements of intervals of time made in the manner indicated above. However, the length of the core is less than the depth of the original undisturbed sediment that was penetrated. A correction must be applied for this shortening which appears to increase regularly from top to bottom of the core. The term distortion is applied here to this shortening. The term compaction will be reserved for the natural phenomenon in the undisturbed sediment.

THE METHOD

The radioelements of the uranium-238 series are not incorporated in equilibrium amounts into the deep-ocean sediments[1] at the time of their deposition. Consequently, the radium content changes as time progresses and therefore as a particular layer becomes buried. The re-establishment of radioactive equilibrium, reflected in the changing radium content, is the basis of a method of measuring time. The method has been adequately described previously (10) and will be discussed only briefly by reference to a typical curve of radium content as a function of depth from measurements in a core of red clay, plotted in Fig. 1.

---

[1] Radioelements are in equilibrium when the rate of growth of each member of a series is equal to its decay rate, and hence the amount present is independent of time. Disturbance of the equilibrium appears to be negligible in the shelf deposits. However, only one core of such a sediment has been studied, and alternative explanations of the constant radium content in this core may be found either in a very high rate of deposition or in the fact that shelf deposits consist of extensively reworked detrital material.

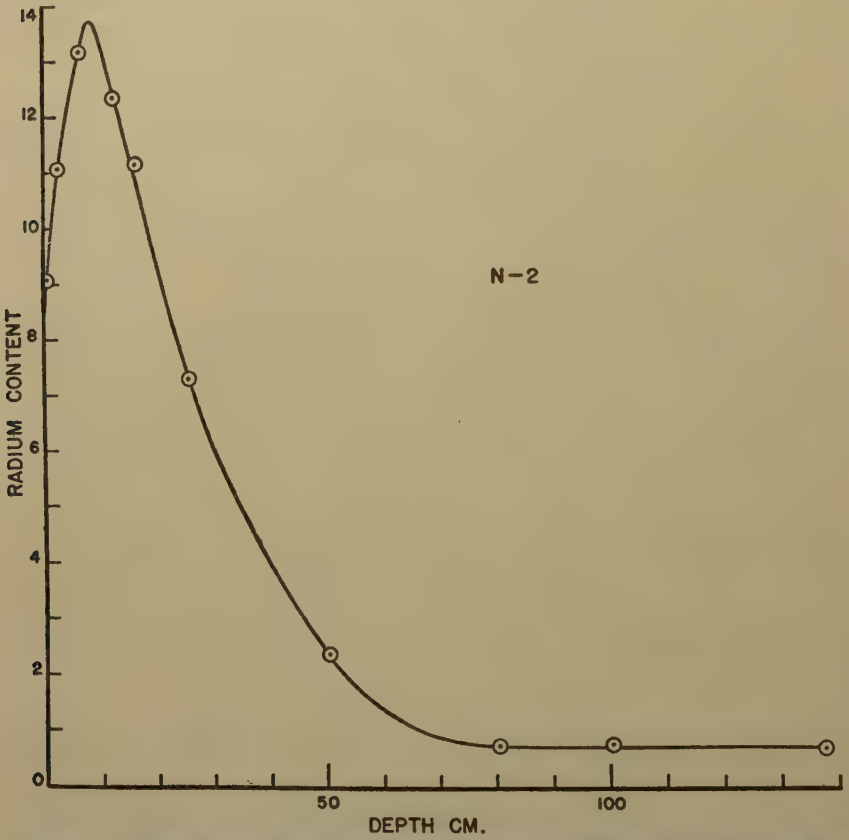


Figure 1. Typical curve of the radium content as a function of depth in a core of deep-sea sediment. The radium content is expressed in units of 10^{-12} g Ra per g of sediment dried at 110° C.

The immediate parent of radium is ionium, which is present in freshly-depositing sediment in excess of that necessary to replenish the radium. Hence the radium at first increases until equilibrium is established between these two elements. The concentration of uranium in deep-ocean sediments is very much less than that necessary to replenish the ionium; consequently the ionium, and with it the radium content, decreases until equilibrium is established between ionium and the small amount of uranium. No further detectable change occurs after this because of the extremely long half-life of uranium; complete equilibrium in the uranium family has been re-established. The half-life of ionium is 82,000 years. Equilibrium is practically re-established in six half-lifes, or roughly 500,000 years. This whole process can be followed by measuring only the radium content. In practice, the decrease in radium becomes so small beyond 450,000 years as to render the method very inaccurate.

In analyzing the data, such as are plotted in Fig. 1, to obtain the age of the sediment at any specific depth it is necessary to know the original ionium and uranium content. These contents can be obtained from the peak value and from the constant radium content in the bottom of the core if the sediment at the bottom is considerably older than 500,000 years, as in core N-2. Many cores, a number of which are much longer than core N-2, were deposited very much more rapidly than this red clay; they show no constant radium content in their lower reaches, but fortunately an accurate knowledge of the uranium content is unnecessary for the first 300,000 years. The uranium content can be measured independently if necessary (14). A general mathematical treatment of the variation of radium concentration with time in a system of radioelements originally present in any haphazard mixture has been published previously (15).

This method has been employed to date a variety of events, such as changes in the lithology and foraminiferal fauna, which have been observed to occur, often repeatedly, in a number of cores from both hemispheres. Such dates seem to constitute the only reliable basis for correlation of these events from place to place. The method will probably find its greatest usefulness in establishing this basis. Nevertheless, it is of value to inquire into the rates of deposition that can be calculated from the dating.

DISTORTION

In securing a core of unconsolidated sediment with a coring tube (3), the sediment is nearly always distorted; that is, the length of the core is less than the depth of sediment penetrated. Because of the thinning of any layer of sediment in the "bulb of compression" established below the bottom of the coring tube, the thickness of the new layers that are forced into the tube as it penetrates deeper into the sediments becomes an increasingly smaller fraction of the *original* thickness of the layers.

The amount of such distortion is not easy to determine, because it depends on a variety of factors such as the type of sediment, the water content, the diameter of the tube, and the velocity of penetration. In order to measure the degree of distortion, it is necessary to core a sediment exhibiting a series of distinctive boundaries which must be measurable in both the core sample and the original undisturbed sediment. Vertical sections of original undisturbed ocean sediments are naturally not readily accessible. Simulated sediments of similar constituency, prepared in tanks, might be used.

Piggot (7) made use of the varved clays of New England to determine the amount of distortion with his gun-coring apparatus (6). A number of the cores studied in the present program of research were secured with this apparatus. It seems correct to assume, for several reasons, that the distortion in coring the varved clay is somewhat greater than would be expected in coring many deep-sea sediments. Hence, application of Piggot's data possibly overcorrects the errors of distortion and thus gives the upper limit for the rates of deposition. With no allowance for distortion, one must obtain the lower limit for the rates. From the data obtained by Piggot, it appears that the distortion at a given depth is a nonlinear function of the fraction of the total core length. It can be seen from Fig. 2 that the effect of distortion is negligible, with respect to the errors in the other determinations, for the upper six-tenths of the total core length. In several cores, the present method of determining the rates of deposition is inapplicable in the lower four-tenths of the core length, because the

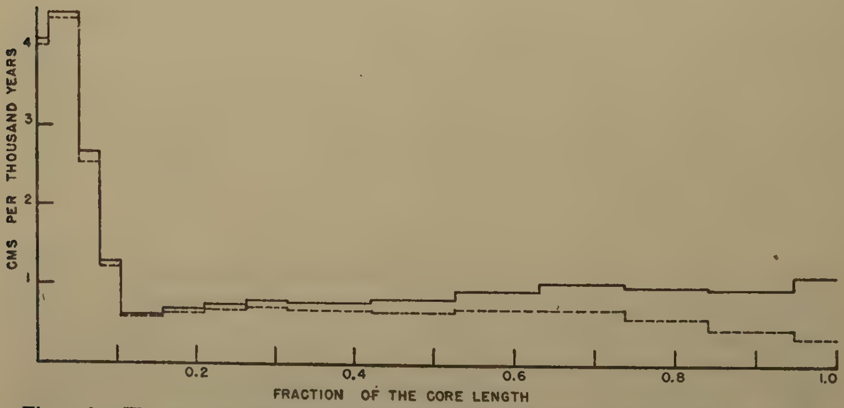


Figure 2. The rate of deposition of globigerina ooze in cm per 1,000 years as a function of the fraction of the core length, illustrating that the error involved in neglecting the distortion becomes appreciable below a depth of 0.6 of the core length. Solid lines—calculated with Piggot's distortion factors. Broken lines—neglecting distortion.

time exceeds 450,000 years; consequently, for such cores the amount of distortion is only in question as far as extrapolation to determine the total time of deposition of the whole core is concerned.

In calculating the following results, Piggot's factors to correct for distortion have been employed except where otherwise stated.

SOURCE AND LOCATION OF MATERIAL

Table I contains the data necessary to evaluate the significance of the rates of deposition. N-numbered cores were secured by Dr. J. L. Hough during the 1946–47 Navy Antarctic Expedition. Samples of core P-259 were provided by Dr. Roger Revelle of Scripps Institution of Oceanography. The remaining P-numbered cores were secured by Dr. C. S. Piggot with his core sampler.

TABLE I. Correlative Data for Cores of Ocean Sediments Studied for Rates of Deposition by Variation of the Radium Content

Core No.	*Type*	*Location*			*Depth*	*Core length*	*Total time of deposition*
		(° ′)	(° ′)		(m)	(cm)	(yrs)
N-3	Glacial marine	68 26 S	179 35 W	Antarctic	3280	228	1,130,000
N-4	Glacial marine	69 12 S	179 34 E	Antarctic	3720	260.5	172,000
N-5	Glacial marine	70 17 S	178 23 W	Antarctic	2980	242.5	460,000
N-2	Red clay	32 21 S	105 55 W	S. E. Pacific	3610	138	1,230,000
P-259	Red clay	30 41 N	121 46 W	N. E. Pacific	4080	246.5	1,080,000
N-1	Red clay-Glob. ooze	08 56 S	92 05 W	Equatorial Pacific	3920	194	990,000
P-135	Globig. ooze	19 18 N	76 48 W	Caribbean	4900	110	21,500
P-136	Globig. ooze	18 38 N	79 12 W	Caribbean	4650	190	111,000
P-137	Globig. ooze	19 14 N	80 20 W	Caribbean	4890	190	301,000
P-126	Foram. marl and Glacial marine	48 38 N	36 01 W	N. Atlantic	4820	282	72,000
P-130		49 40 N	28 29 W	N. Atlantic	3740	277	24,000
P-124	Calc. blue mud	46 03 N	43 23 W	N. Atlantic	4700	281	12,000

RATES OF DEPOSITION

The rates of deposition for the past 450,000 years are shown in Figs. 3 and 4. Few general conclusions can be reached from the study of such a small number of cores. Red clay from the North and South Pacific and glacial marine deposits of the Antarctic Ocean show (Fig. 3) no immediately apparent change in the rate of deposition at the time of the last interglacial stage in the northern hemisphere. The rate of deposition was generally low at the time of the last glacial stage in the northern hemisphere. The only notable changes in the rates of deposition during the past 450,000 years appear to have occurred at the time of the middle interglacial stage, the Yarmouth of North America or the Mindel-Riss of Europe. During this middle interglacial stage some of the cores show an increased rate of deposition, occasionally almost as rapid as that at present.

Rates of deposition and their variation with the climatological conditions must be controlled by complex factors, often opposing in their effects. These factors have decided local characteristics, as is evident from the high rates in cores P-124 and P-130. These high rates, as previously discussed (10) and given here in Table II, are undoubtedly due to the locations of the sediments at the foot of the Continental Shelf off Newfoundland and at the foot of the east wall of the Faraday Hills of the mid-Atlantic Ridge. The high rate of deposition of the glacial marine sediment at the site of core N-4 at one end of a trough south of Scott Island Bank, compared with the rates at adjacent sites N-3 and N-5, is possibly due to a similar condition.

Deposition of globigerina ooze in the Caribbean Sea was also slowest during the last glacial stage in the northern hemisphere, as shown in Fig. 4 (cores P-135, 136, 137). There appears to be some evidence that the increase to the present high rates of deposition of globigerina ooze was delayed by several thousand years as compared with the red clay of the Pacific and with the glacial marine deposits of the Antarctic. In fact, there was little or no increase in the rate of deposition of globigerina ooze in the Caribbean Sea until the beginning of the climatic optimum of postglacial time about 7,000 years ago. This delay had been noted previously in the foraminiferal marls of the North Atlantic (10) and is attributed to the fact that considerable time is necessary to re-establish the abundance of organic life following a prolonged period of cold surface waters.

An interesting phenomenon is observed in cores P-135 and P-136 from the Caribbean Sea. At the site of both of these cores there was comparatively instantaneous deposition of 70 to 80 cm of sediment, as shown in Fig. 4; and in each case, this rapidly-deposited globigerina ooze was characterized by a 2 cm thick zone of practically pure, uncemented foraminiferal shells at a depth of about 66 cm in both cores. Instantaneous deposition of any such depth as 70 to 80 cm of sediment is immediately apparent from the curves of radium content. Thus, the typical curve of Fig. 1 will be sharply broken, with the lower portion displaced to greater depths, the two parts being joined by a horizontal straight line (9). The break occurs at 21,500 years in core P-135 and at 21,000 years in core P-136, which is close enough to be considered contemporaneous. In view of the probable error of the

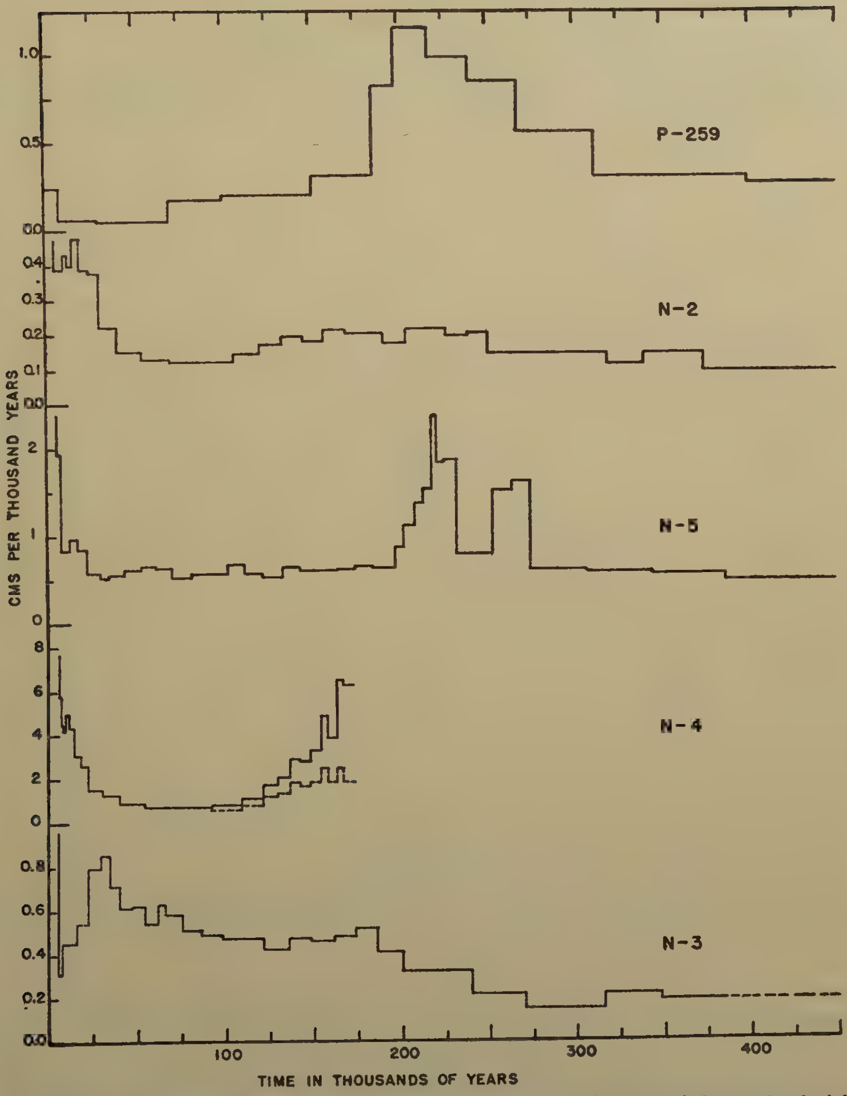


Figure 3. Variations of the rate of deposition of Pacific red clay and Antarctic glacial marine sediments in the past. Red clay—P-259 and N-2. Glacial marine sediments—N-3, N-4 and N-5. N-4—broken lines—no correction for distortion; the distortion factors very probably overcorrect the rate of deposition at the bottom of this core.

measurements, it cannot be stated that the deposition was truly instantaneous, but it must have occurred in less than 1,000 years. The rapidly-deposited sediment in both cores consists largely of foraminiferal remains and thus this observation may indicate a major catastrophe to organic life in the Caribbean Sea at the height of the marine equivalent of the last glacial substage. There are alternative explanations, however, and the same phenomenon may be caused by a major tectonic event. Core P-137, also from the Caribbean Sea, shows no such break in the deposition. Core P-136 was secured from the center of an extensive plateau forming the bottom of the Cayman Trough south of the Bartlett Deep. Core P-135 was secured from the same plateau at the eastern end of the Cayman Trough just south of the Oriente Deep. Core P-137 was obtained at 4,890 meters from the steep slope of the Bartlett Deep between Little Cayman Island and the

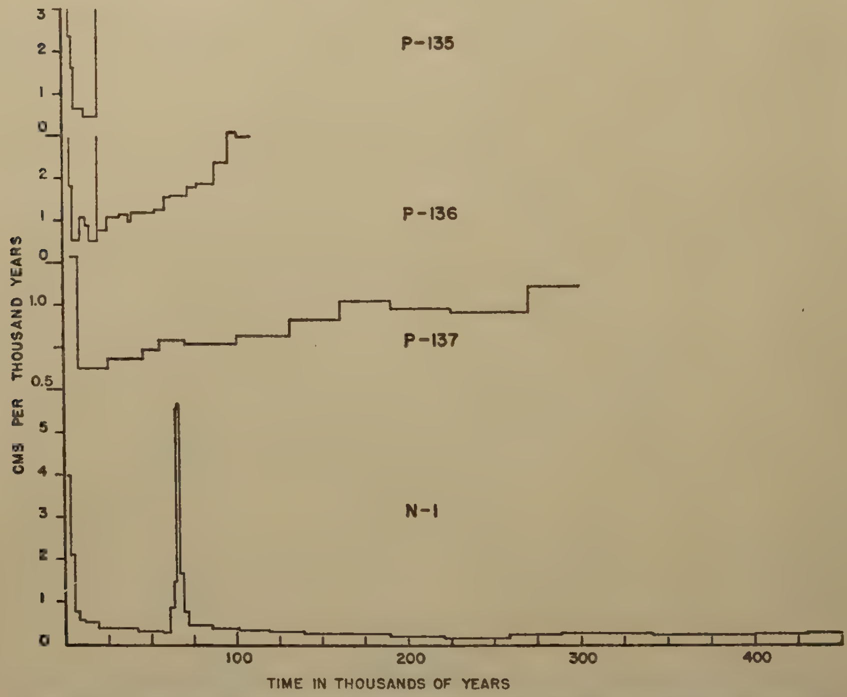


Figure 4. Variations of the rate of deposition of Caribbean Sea globigerina ooze and of Pacific globigerina and red clay mixture (N-1) in the past. The vertical lines at 21,000 years in P-135 and P-136 indicate practically instantaneous deposition of a long section of each of these cores (see text).

bottom of the Deep at 7,000 meters. However, there is no evidence in core P-137 that the steep walls of the Trough exhibit the counterpart of instantaneous deposition on the floor of the Trough. The loss of any appreciable section of sediment, unless it occurred immediately following deposition, should be characterized by a vertical break in the curve of Fig. 1.

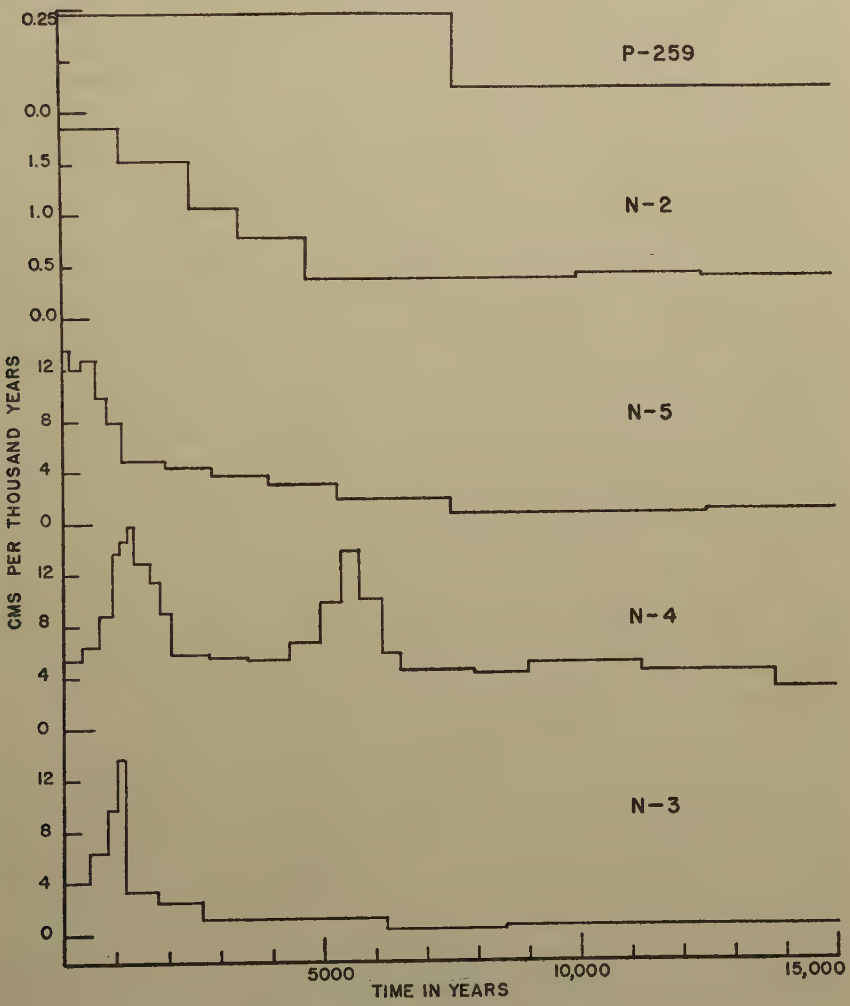


Figure 5. Showing the detail of Fig. 3 for the past 15,000 years.

An unpublished study of the temporal correlation of the relative abundance of the foraminiferal types in these three cores, as determined by Cushman (2), verifies the extraneous nature of the rapidly-deposited sediment on the plateau of the Trough, in the stratigraphic column.

An outstanding feature of the deposition of nearly all deep-sea sediments is the high rate in postglacial time, particularly since the beginning of the first climatic optimum about 7,000 years ago. This is not due to the compaction of the sediments, as can be demonstrated by determining the weight of solid material deposited per unit area per year (10), or by the centimeters of pore-free sediment per 1,000 years as given in Table II. Although some decrease in water content with depth in a core has been reported (13), many cores show no systematic change in the water content. In any case, the amount of compaction corresponding to any observed change in the water content is far too small to account for the high rates of deposition at present compared with the past. Moreover, there is correspondence in the rates of deposition as a function of time in a suite of cores wherein the sediment may be 10,000 years old, a few centimeters below the top in one and tens of centimeters below the top in another (10). For this same reason, it does not seem possible to consider the high rates near the surface as spurious and as caused by leaching of the radioelements with a displacement of the peak radium content (Fig. 1) to greater depths. Under such circumstances it would be impossible to derive the same date of 12,000 years in several cores for the recent commencement of deposition of volcanic glass shards in the North Atlantic, when this boundary occurs anywhere from 20 to 280 cm below the top (1).

Thus it appears necessary to accept the generally high rates of deposition in postglacial time. Figs. 5 and 6 show the variation of the rate of deposition in the past 15,000 years. The rate of sedimentation of red clay (Fig. 5) appears not to have increased recently to the same extent as the rates for glacial marine sediment (Fig. 5) or globigerina ooze (Fig. 6) and foraminiferal marl (10).

The rates of deposition of the glacial marine sediments of the Antarctic seem to bear a relation to the principal climatic optimum which reached its peak 4,000 to 6,000 years ago, and to the secondary climatic optimum of 950 to 1,550 years ago. This is particularly evident in core N-4 where the analysis is more detailed in the immediate past because of the length of the core representing the last 15,000 years. Radium determinations would have to be spaced much closer together to obtain the same detail in cores N-3 and N-5. High rates of deposition at these times might be explained by a withdrawal polewards of the limits of the drift ice. These limits at present are appreciably north of the location of these sediments. Glacial debris should be most abundant near the limits of the drift ice. Extending this hypothesis to glacial stages, it might be supposed that the very low rates of deposition of the glacial marine sediments during these stages are due to a permanently frozen Antarctic Ocean reaching north at least to Scott Island Bank at about Lat. 67° S. During these stages the sediment should be relatively free of glacial pebbles and fragments and there is some evidence from preliminary studies that this was the case at the height of the glacial stages.

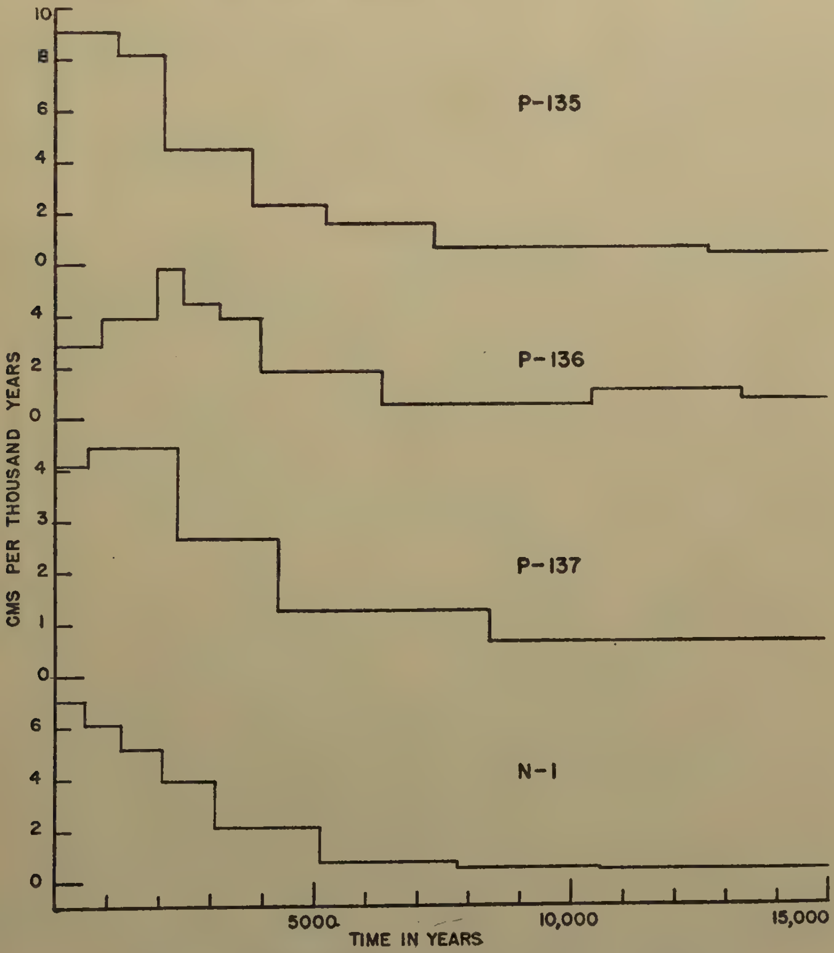


Figure 6. Showing the detail of Fig. 4 for the past 15,000 years.

Such an explanation supposes that climatic changes were roughly contemporaneous over the whole world. It is hoped that the still-unsolved question of simultaneous or alternating climatic changes and glaciations in the two hemispheres can be answered by studies of these cores, but it is certainly not possible to draw any conclusion from the rates of deposition. An answer to this question should be obtainable by studying the lithology and foraminifera of the cores from the southern hemisphere and correlating such studies with those already performed for sediments of the northern hemisphere on the basis of time. These studies are under way.

The globigerina ooze of the Caribbean Sea was not deposited at any excessive rate during the first climatic optimum, although this period seems to have inaugurated the general increase in the recent rate of deposition of this ooze. This conforms to the previously expressed idea that changes in the population of the surface organic life of the ocean occur only slowly.

In core N-1 from the equatorial Pacific, foraminifera are present throughout the core, but there are only a few shallow zones of pure globigerina ooze. For the most part globigerina ooze is mixed with red clay. In postglacial time (northern hemisphere) the sediment was largely red clay. During the last glacial stage (northern hemisphere) red clay alternated with globigerina ooze. These alternations appear to correlate with the warm and cold substages of the last glacial stage in the sediments of the North Atlantic and Caribbean Sea. The overall low rate of deposition of the sediment at the site of core N-1 is not surprising, therefore, because the conditions are bordering on those which lead to the deposition of red clay.

COMPARISON OF RESULTS WITH PREVIOUS ESTIMATES

The rate of deposition of marine sediments, even for those of similar composition, is so variable from place to place and from time to time that average values have little or no significance.

Schott's methods for the equatorial Atlantic, not corrected for the pore space (water content), gave him estimates for the past 20,000 years of between 2.13 and 0.53 cm per 1,000 years for globigerina ooze and between 1.33 and less than 0.5 cm per 1,000 years for red clay. The average values over the past 15,000 years, as determined by the radioactive method, are shown in Table II. They lie within the limits of Schott's estimates.

TABLE II. Average Rates of Deposition of Deep-sea Sediments in Cm per 1,000 Years for the Sediment in the Original State (column A) and for the Sediment in a Pore-Free State (column B)

Core No.	*Type*	*Per cent solid by volume**	*Time interval in thousands of years*							
			0 to 15		15 to 200		200 to 300		300 to 450	
			A	B	A	B	A	B	A	B
N-3	GM	30	1.3	0.39	0.50	0.15	0.23	0.07	0.22	0.07
N-4	GM	31	5.9	1.8	1.9†	0.59	—	—	—	—
N-5	GM	33	2.5	0.83	0.61	0.20	1.1	0.36	0.50	0.17
N-2	RC	31	0.68	0.21	0.18	0.06	0.16	0.05	0.12	0.04
P-259	RC	24	0.16	0.04	0.21	0.05	0.81	0.19	0.27	0.06
N-1	RC-GO	17	1.8	0.31	0.38	0.06	0.21	0.04	0.25	0.04
P-135	GO	42	2.5	1.1	—	—	—	—	—	—
P-136	GO	36	1.7	0.61	1.6†	0.58	—	—	—	—
P-137	GO	44	1.7	0.75	0.82	0.36	1.0	0.44	—	—
P-126	FM-GM	33	1.8	0.59	6.7†‡	2.2	—	—	—	—
P-130	FM-GM	26	15.0	3.9	20.0†‡	5.2	—	—	—	—
P-124	CBM	25	33.7†	8.4	—	—	—	—	—	—

† The bottom of the core is younger than the upper limit of the time bracket (see Table I).

‡ This interval includes the glacial marine deposits of the equivalent of the Wisconsin glacial substages in the North Atlantic.

* The per cent of solid by volume was determined from the water content by weight and an assumed grain density of 2.6 gr. per cc.

The average rate of deposition is given in Table II over somewhat arbitrary intervals of time so chosen that, on the whole, the rate of deposition remained roughly constant in any one interval. The period from 15,000 to 200,000 years would roughly cover the last glacial and interglacial stages and the third glacial stage. The period from 200,000 to 300,000 years would cover the long middle interglacial stage and the period from 300,000 to 450,000 years would cover the early part of the middle interglacial stage and possibly part of the second glacial stage. These dates are for the northern hemisphere (10).

Insofar as one can judge from the few cores studied, there is some evidence of increased rate of deposition in the long middle interglacial stage, but, with one exception, the deposition was never so rapid as at present. There is no evidence that the last interglacial stage had any profound effect on the rate of deposition of deep-sea sediments. This leads one to suspect that postglacial rates of deposition are far from representing the average rate of deposition throughout geological time.

For a comparison of the measurements reported here, with the estimates based on supply methods, it is necessary to reduce the values to a pore-free state. This can be accomplished with a knowledge of the water content by volume which was determined for each sample. The average per cent of solid material by volume is shown in Table II for each core.

Kuenen (4), from a consideration of the extent of erosion since the Cambrian (500 million years), estimated 0.1 cm per 1,000 years for red clay and 0.2 cm per 1,000 years for globigerina ooze. The values in Table II for the warm periods of the middle interglacial stage and postglacial time give an average of 0.11 cm per 1,000 years for red clay. If the average over geological time is nearer to the mean over the whole of the last half-a-million years, a figure of 0.07 cm per 1,000 years would be more appropriate. Kuenen's estimate for globigerina ooze appears to be low on the basis of the present measurements in the Caribbean Sea, but this is hardly a fair comparison. However, it is definitely low for the foraminiferal marl of the North Atlantic (P-126).

Other estimates based on the supply methods are compared with Kuenen's values in *The Oceans* (13) and need not be discussed here.

SUMMARY

A study of the variation of the radium content during the period of re-establishment of radioactive equilibrium in the buried deposits of deep-sea sediments provides a method of dating the record of past events in the ocean bottom. The results of such studies can be readily applied to determinations of the rate of deposition provided that a knowledge of the distortion in securing core samples of the deep-sea sediments is available.

Application of the method of discerning rates of deposition is not limited, as was the application of earlier methods, to sediments deposited during postglacial time; it is possible to study the variation of the rate of deposition in the past as far back as the method of dating is applicable, *i. e.*, for about half a million years. Rates of deposition as a function of time are reported here for red clay, globigerina ooze, foraminiferal marl, glacial marine deposit, and calcareous blue mud, from the Antarctic Ocean to the North Atlantic.

Outstanding features of these determinations are as follows: Deposition of practically all of the sediments is more rapid at present than during the past half million years. The repeated climatological changes of the ice age did not have a particularly noticeable effect on the rate of deposition in general. Only during the long middle interglacial stage did the rates tend to be somewhat higher than the low rates generally prevailing in the past half million years. The lowest rates of deposition are associated with the last glacial stage and the early period of the middle interglacial stage or possibly the end of the

second glacial stage. Locally, there are often interesting short-period changes in the rate of deposition which appear to be caused by climatological changes.

However, the amount of detail in an analysis of the rates of deposition varies greatly: in cores of equal length, far more detail can be obtained in one where the sediment was deposited in 100,000 years than in one where the sediment required a million years for deposition.

Average rates of deposition, insofar as they are comparable, are in good agreement with previous estimates by the *stratigraphic* and *supply* methods.

REFERENCES

1. Bradley, W. H., *et al.*
   1940. Geology and biology of North Atlantic deep-sea cores. Prof. Pap. U. S. geol. Surv., 196-A: I–XV + 1–50.
2. Cushman, J. A.
   1941. A study of the Foraminifera contained in cores from the Bartlett Deep. Amer. J. Sci., *239:* 128–147.
3. Hough, J. L.
   1939. Bottom-sampling apparatus; In, Recent Marine Sediments, A Symposium. Amer. Ass. Petrol. Geol. (pp. 631–664). 736 pp.
4. Kuenen, Ph. H.
   1941. Geochemical calculations concerning the total mass of sediments in the earth. Amer. J. Sci., *239:* 161–190.
5. Lohmann, H.
   1909. Plankton-Ablagerungen am Boden der Tiefsee. Schr. naturw. Ver. Schl.-Holst., *14:* 399–402.
6. Piggot, C. S.
   1936. Apparatus to secure core samples from the ocean-bottom. Bull. geol. Soc. Amer., *47:* 675–684.
7. 1941. Factors involved in submarine core sampling. Bull. geol. Soc. Amer., *52:* 1513–1523.
8. Piggot, C. S. and W. D. Urry
   1941. Radioactivity of ocean sediments. III. Radioactive relations in ocean water and bottom sediment. Amer. J. Sci., *239:* 81–91.
9. 1942. Radioactivity of ocean sediments. IV. The radium content of sediments of the Cayman Trough. Amer. J. Sci., *240:* 1–12.
10. 1942. Time relations in ocean sediments. Bull. geol. Soc. Amer., *53:* 1187–1210.
11. Revelle, Roger and F. P. Shepard
    1939. Sediments off the California coast; In, Recent Marine Sediments, A Symposium. Amer. Ass. Petrol. Geol. (pp. 245–282). 736 pp.
12. Schott, Wolfgang
    1935. Die Sedimente des äquatorialen Atlantischen Ozeans. B. Die Foraminiferen in dem äquatorialen Teil des Atlantischen Ozeans. Wiss. Ergebn. dtsch. Atlant. Exped. *Meteor*, 1925–27, Bd. 3, Teil 3, Lief. 1B: 43–134.

13. Sverdrup, H. U., M. W. Johnson and R. H. Fleming
    1942. The Oceans. Prentice Hall, Inc., New York. 1087 pp.

14. Urry, W. D.
    1941. The radioactive determination of small amounts of uranium. Amer. J. Sci., *239:* 191–203.

15. 1942. The radio-elements in non-equilibrium systems. Amer. J. Sci., *240:* 426–436.

16. Radioactivity of ocean sediments. VI. Concentrations of the radio-elements in marine sediments of the southern hemisphere. Amer. J. Sci. (In Press.)

17. Urry, W. D. and C. S. Piggot
    1942. Radioactivity of ocean sediments. V. Concentrations of the radio-elements and their significance in red clay. Amer. J. Sci., *240:* 93–103.

RADIUM ANALYSES OF MARINE SEDIMENTS IN NORTHERN PACIFIC AND ADJACENT WATERS

By

C. L. UTTERBACK AND L. A. SANDERMAN
University of Washington
Seattle, Washihgton

Although systematic soundings of the ocean bottom were initiated approximately 100 years ago, and although the United States Coast and Geodetic Survey took hundreds of ocean bottom samples beginning about 75 years ago, it was not until the ship CARNEGIE collected a number of bottom samples at great depths that information began to be available on the radioactive properties of deep ocean sediments. An analysis of some of the CARNEGIE bottom samples by Joly (6) showed that they had a considerable range of radium content and that some of them contained from 16 to 20 times as much radium per gram of material as was found in land rocks. This relatively high concentration of radium in deep sea sediment was also found by Pettersson (7) in analyses of samples collected by the PRINCESS ALICE II.

These early results, verified later (8, p. 229), showed that deep ocean sediments have a radium content as high as $22 \cdot 10^{-12}$ g of radium per gram of material. Undoubtedly this high value was an important incentive for further investigation, since continental granites have only about $1.6 \cdot 10^{-12}$ g of radium per gram of material. The basic igneous land rocks and the sedimentary igneous land rocks have even less.

Sea water from various localities and depths have been analyzed. The measurements of Evans, *et al.* (3, p. 247), which were made on 13 samples, yielded an average of $0.8 \cdot 10^{-16}$ g of radium per cc of water. These samples from the Pacific, Lat. 32° 32′ N., Long. 117° 22′ W., were obtained at depths which varied from the surface to 1,300 meters. Also, Foyn, *et al.* (4, p. 27) found for 23 samples an average value of $0.86 \cdot 10^{-16}$ g of radium per cc of water.

A systematic survey of the radium content of sediments in the waters of Vancouver Island were made by Utterback and Sanderman (13, p. 187). These samples were from locations such that one would conclude that they were deposited by drainage from Vancouver Island and the coast of British Columbia. This territory is composed largely of schist and quartzite with a small amount of dolorite. The average radium content of samples from 11 stations was $0.24 \cdot 10^{-12}$ g of radium per gram. In Puget Sound, samples from five stations showed an average content of $0.37 \cdot 10^{-12}$ g of radium per gram. The surrounding territory is made up largely of glacial drift with a radium content comparable with the average of the five stations.

Sanderman and Utterback (10, pp. 132, 137), in a more extensive survey, obtained many samples of sediments from the Arctic Ocean, Norton Sound, Bering Sea, the Alaska Peninsula, southern Alaska, the Canadian Coast, and Queen Charlotte Islands. The concentration in these samples varied from $0.24 \cdot 10^{-12}$ to $1.57 \cdot 10^{-12}$ g of radium per gram, several samples showing relatively high values. Those of high radium concentration were all taken from glacier-fed inlets, and the average was $1.31 \cdot 10^{-12}$ g of radium per gram. Several samples taken from the region of the mouth of the Skeena River showed an average of $0.63 \cdot 10^{-12}$ g of radium per gram, while representative rocks furnished by the Chief Geologist of the Department of Mines and Resources of Canada had an average radium content of $0.64 \cdot 10^{-12}$ g of radium per gram. Obviously this does not permit of a general conclusion, but it seems to indicate a radium content of sediments comparable to that of adjacent land areas or drainage areas. In this study the highest average was from the region of the southern Alaska Coast where most of the rivers rise from the glacial slopes of the near-by granitic mountains. The lowest average is from the volcanic Alaska Peninsula.

The bottom samples referred to above were collected by means of a clam shell sampler or a short length of pipe. But in any case the bottom was sampled to a depth of a few centimeters. The results are in confirmation of the hypothesis that the radium content of inshore and nearshore bottom samples is due to the deposition from the surrounding land areas.

In addition to the samples mentioned above, five grab samples from the Northeast Pacific are reported by Piggott (8, p. 231) to have an average radium content of $7.6 \cdot 10^{-12}$ g of radium per gram. A sample taken from the bottom of water intermediate between the deep ocean and the near land was collected off the coast of California. It was analyzed by Evans (2) and was found to have $3.2 \cdot 10^{-12}$ g of radium per gram. Piggott (9, p. 675) obtained eleven cores between the Grand Banks of Newfoundland and the Continental Shelf southwest of Ireland. In these cores the red clay varied from $3 \cdot 10^{-12}$ to $22 \cdot 10^{-12}$ g of radium per gram, the globigerina ooze from $3 \cdot 10^{-12}$ to $7 \cdot 10^{-12}$ g of radium per gram, and the blue mud from $1 \cdot 10^{-12}$ to $3 \cdot 10^{-12}$ g of radium per gram. The higher concentrations were in samples taken at locations far from land masses and under deep water.

The high concentration of radium in deep sea deposits has led to some speculation about the process by which radium is deposited. Four possible processes have been considered. These are chemical precipitation, biological concentration, absorption of radium from settling particles, and submarine volcanic activity. However, the mode of deposition of radium in ocean water is not known, probably because the biological, chemical, and physical processes and changes in the deep ocean waters have not been determined with sufficient completeness.

Experimental data are available to test some of the processes suggested to account for the high radium content of deep sea sediment.

It is well known that marine organisms are capable of concentrating some of the constituents of ocean water. Berget (1) suggests that perhaps the radium content of deep sea sediments is due to organic accumulation. Evans, *et al.* (3, p. 256) determined the radium content of samples of plankton taken near La Jolla, and samples of kelp from Mission Beach, California. Using the weight of the samples before they were reduced to ash, they concluded that the radium concentration in living plankton and kelp is about 100 times as great as that of the water in which they live. However, when these samples were reduced to ash, the plankton had a concentration of $0.46 \cdot 10^{-12}$ g of radium per gram, and the kelp a concentration of $0.21 \cdot 10^{-12}$ g of radium per gram.

There is some evidence relative to the possible chemical extraction of radium. Foyn, *et al.* (4, p. 31) have compared the effect of calcium extraction by marine organisms and have compared the radium to calcium ratio in calcareous shells. They concluded that the rate of extraction of radium is no greater than the rate of extraction of calcium from sea water. Since the rate of extraction of calcium changes the calcium content of water but little, then any large rate of extraction of radium must be of some other source. However, since the precipitation of iron hydroxide is a method for precipitating ionium from dilute solutions, such a process might conceivably occur. It has been observed by Thompson, *et al.* (11, p. 288) that the concentration of ferric iron in ocean water is less than it should be if all the iron received were held in solution. The precipitation of uranium would be inhibited by the presence of the carbonate ion. This could contribute to the means by which the ionium-uranium unbalance in the sea water and sediment occurs. However, the question arose as to whether the extraction process concerned radium or a parent substance. Tests were made on the uranium content of sea water and on samples of high radium concentration. Hernegger and Karlik (5) found that sea water had an average of $1.3 \cdot 10^{-9}$ g of uranium per gram of water, but the average of many observers in many determinations is $0.8 \cdot 10^{-16}$ g of radium per gram of water. If these averages are assumed to be correct, then the average amount of radium per gram of water could be supported by $0.24 \cdot 10^{-9}$ g of uranium per gram of water, so the sea water would contain five times as much uranium as necessary to maintain the radium content. It was found by Urry (12, pp. 199–203) that a bottom sample of high radium content had about one-fourth as much uranium as necessary to support its radium content. However, the uranium concentration in these bottom samples would not be the determining factor in the radium concentration. As Correns (1a) has estimated, the rate of deposition in the deep ocean is about one centimeter per 1,000 years, and therefore the uranium and radium may not be in equilibrium. Whatever the process, it must be efficient, since high concentration bottom samples contain tens of thousands of times as much radium per unit weight as does the ocean water.

The experimental data and the analyses up to 1941 show a definite pattern of radium concentration in ocean sediments. The concentrations varied with high values under deep water far from land, lower values in the samples taken off the coast, and concentrations comparable with surrounding land rocks in samples taken under inshore waters. In view of the fact that a systematic survey had been made of the radium contents of bottom samples from the surface of the bottom in the region extending from the northern tip of Alaska to the southern end of Puget Sound, it seemed desirable in this new survey over the same area to obtain samples of the sediment to as great a depth as possible below the ocean floor. To this end, a core sampler designed by Piggott (9, p. 675) was provided, and with it a number of cores up to ten feet in length were obtained in selected regions in the inshore waters of Puget Sound and the San Juan Archipelago. While the analyses of the cores have not been completed (because of interruption during the war), some positive evidence has been obtained concerning the radium concentration of sediments under the inshore water.

One of the regions included the area from the mouth of the Skagit River through Skagit Bay; the material brought down by the river is deposited along the length of the Bay. Another was Hood Canal, which is a long narrow body of water extending from Puget Sound into the Olympic Peninsula. The third region was East Sound in the San Juan Archipelago; this was selected because it is especially rich in plankton.

SKAGIT BAY

The region drained by the Skagit River and its tributaries lies east of Puget Sound and extends into the Cascade mountains. A geological report by Weaver (14) describes most of the region, the greater part of which is composed of sandstone, slate, and shale. One broad granitic area extends from Mt. Shukson on the west to the north-south stretch of the Skagit River on the east. A small region of basic andesitic lavas is found in the Mt. Baker district and also in the higher peaks of the Cascades. Skagit Bay is very shallow at the mouth of the river and deepens to about 75 fathoms down the Bay. Four stations were occupied and cores taken at each.

Station I was just off the river mouth where the water was 10 fathoms deep. The analysis of this core showed a radium content in grams of radium per gram of $0.60 \cdot 10^{-12}$ in the top three inches, $0.58 \cdot 10^{-12}$ in a sample from a depth of 27 inches in the core, $0.71 \cdot 10^{-12}$ from 51 inches, and $0.67 \cdot 10^{-12}$ from 60 inches.

Another core in Skagit Bay was obtained at Station II where the water was 15 fathoms deep and in a region where the water began to deepen rapidly. At distances down the core of 11, 35, 59, and 83 inches, the radium content was in grams of radium per gram $0.58 \cdot 10^{-12}$, $0.58 \cdot 10^{-12}$, $0.64 \cdot 10^{-12}$, and $0.55 \cdot 10^{-12}$ respectively.

The third core was from further down the Bay at Station III where the depth of water was 45 fathoms. At depths in the core of 3, 27, 51, and 75 inches, the radium concentration in grams of radium per gram was $0.52 \cdot 10^{-12}$, $0.50 \cdot 10^{-12}$, $0.60 \cdot 10^{-12}$, and $0.57 \cdot 10^{-12}$.

At Station IV the water was 75 fathoms and the core was nine feet in length. While the variation in radium concentration (with respect to depth in the core) was about the same for the first three stations, it differed considerably at the fourth station. In the cores mentioned above, the maximum radium concentration was found at about five feet below the surface of the core, after which it decreased. At the fourth station the concentration rose from a value of $0.35 \cdot 10^{-12}$ g of radium per gram at 3 inches to a value of $0.62 \cdot 10^{-12}$ g of radium per gram at 27 inches. The concentration was quite constant to about 60 inches, then rose to a maximum of $0.71 \cdot 10^{-12}$ g of radium per gram at 75 inches and gradually decreased to $0.43 \cdot 10^{-12}$ g of radium per gram at 106 inches.

Station II was 3 miles from Station I, Station III 5 miles further, and Station IV 7.5 miles further down the Bay, making a total distance of 15.5 miles from the flat at the mouth of the river. The average of 23 samples from four cores was $0.6 \cdot 10^{-12}$ g of radium per gram.

Although the results of the analyses show that the radium content rises to a maximum at some depth below the surface of the core, further analyses will be necessary to attempt an explanation of the rate of deposition of radium which will account for the experimental results.

In so far as the analyses have proceeded, a consideration of the average radium content of the cores and the nature of the drainage area constitutes further evidence that the radium concentration of inshore sediments, even at considerable depths, is due to a deposition of sediment from the surrounding land areas.

HOOD CANAL

The second region selected was Hood Canal, which is a narrow body of water whose total length is approximately 60 miles. It empties into Puget Sound at the north end, while the south end, which extends into the Olympic Peninsula, is closed. There are numerous rivers and streams flowing in from the Olympic mountains on the west and a few streams from the land area on the east. The stations selected for sampling, numbered from I to V inclusive, are located, with respect to the junction of the Canal with Puget Sound, at distances of 24, 29, 45, 49, and 53 miles respectively. The respective water depths were 74, 79, 74, 64, and 41 fathoms.

The lengths of the cores were as follows: 7.5 feet at Station I, 8 feet at Station II, 10 feet at Stations III and IV, and 6 feet at Station V. The cores showed no stratification at Stations I and V, very little at Stations II and IV, but considerable stratification at Station III.

The core taken at Station I had a uniform radium concentration throughout its length with an average value of $0.58 \cdot 10^{-12}$ g of radium per gram. The core from Station II also showed a uniform distribution with an average value of $0.61 \cdot 10^{-12}$ g of radium per gram. The core from Station III had a concentration of $0.86 \cdot 10^{-12}$ g of radium per gram in the upper three inches of its length; in the next five feet the concentration was fairly constant at $0.79 \cdot 10^{-12}$ g of radium per gram; the concentration decreased to $0.38 \cdot 10^{-12}$ g of radium per gram at 6 feet 3 inches and remained fairly constant, with a value of $0.44 \cdot 10^{-12}$ g of radium per gram at the bottom of the core. The core from Station IV had constant concentration for the first 8 feet of its length with an average value of $0.58 \cdot 10^{-12}$ g of radium per gram, but the concentration at 9 feet 9 inches was $0.33 \cdot 10^{-12}$ g of radium per gram. The radium concentration along the core taken at Station V was nearly constant with an average value of $0.36 \cdot 10^{-12}$ g of radium per gram.

The study of the cores from Hood Canal is incomplete and provides but little data from which positive conclusions as to rate of deposition of radium might be based. However, some conclusions from the preliminary results of analyses are suggested. The geological structure of the drainage areas of the rivers flowing into Hood Canal has been reported by Weaver (14). Most of the west side of the Canal is composed of basalts extending from the shore of the Canal into the Olympic mountains. Some sandstone appears along the west shore from the location of Station II to Station III. Extending from Station III to Station V is a band of glacial material covering the underlying basalt. The materials exposed on the east side of the Canal are glacial sand, gravel, and clay.

The analyses of the cores show a range in the radium concentration extending from $0.33 \cdot 10^{-12}$ to $0.86 \cdot 10^{-12}$ g of radium per gram. This range is quite comparable to the range extending from the mean value of basic igneous rock, such as basalt, to a value between that of the intermediate and acidic igeneous rock, such as hornblende and granite. For the most part the cores from Stations I, II, and IV had radium concentrations comparable to that of intermediate and sedimentary igneous rock.

It is obvious that the sediment in Hood Canal is well mixed by river and tidal currents on entering the Canal. The rivers from the west side of the Canal rise in the Olympics at elevations of approximately 5,000 feet. They reach the Canal with relatively high currents and have a greater effect on the distribution of the sedimentary material than do the small rivers entering the Canal from the east, which rise from elevations of about 1,000 feet and have relatively low currents. The number of rivers entering from the west is much greater than those entering from the east. Also, there is a mixing of the materials along the axis of the Canal by tidal action.

EAST SOUND

This Sound is a long narrow body of water nearly surrounded by Orcas Island, the largest island in the San Juan Archipelago. It is approximately seven and one-half miles long and one to one and one-half miles wide. This body of water is of particular interest in as much as it contains a high concentration of plankton.

A core nine feet long was obtained at a station four miles from the mouth of the Sound at a water depth of 18 fathoms. The radium concentration was determined at one-foot intervals beginning at the surface. The samples from the first seven feet of the core, with the exception of those at the two- and five-foot depths, had a radium concentration varying between the values of $0.66 \cdot 10^{-12}$ and $0.62 \cdot 10^{-12}$ g of radium per gram. The samples taken from the two- and five-foot depths had a radium concentration of $0.52 \cdot 10^{-12}$ and $0.53 \cdot 10^{-12}$ g per gram respectively. The bottom foot of the core exhibited a decrease in radium concentration to an average value of $0.57 \cdot 10^{-12}$ g of radium per gram.

Since the average radium concentration of the East Sound core is essentially the same as that of the cores from the Hood Canal and Skagit regions which are not rich in plankton, it seems that the biological extraction of radium must not be of prime importance in the radium deposition process.

The above analyses and the conclusions from the data so far obtained constitute a preliminary report on the cores obtained in Skagit Bay, Hood Canal, and East Sound. The completion of the mineralogical and mechanical analyses of the cores may modify or supplement the above conclusions.

The authors acknowledge the assistance of Dr. R. E. Wilson, Dr. D. P. Burcham, and Mr. Reinertsen in the analyses of the cores.

SUMMARY

Several hundred samples of marine sediments from the surface of the ocean bottom have been analyzed for radium content. These bottom samples were from the region extending from the northern tip of Alaska to the southern end of Puget Sound, including the Arctic Ocean, Bering Sea, the Alaska Peninsula, southern Alaska, Queen Charlotte Islands, Vancouver Island, and the San Juan Archipelago. The radium content varied from $0.24 \cdot 10^{-12}$ to $1.57 \cdot 10^{-12}$ g of radium per gram. The samples of high radium content were all from glacier-fed inlets.

The analysis of 23 samples from four cores gave an average value of $0.6 \cdot 10^{-12}$ g of radium per gram. These cores were from selected stations in the Puget Sound area and in the San Juan Archipelago.

The data show a definite pattern of radium concentration in ocean sediments. The radium content of samples taken off the coast or from inshore waters is comparable to that of rocks in near-by land areas or of rocks in adjacent drainage areas.

BIBLIOGRAPHY

1. Berget, Alphonse
 1930. Leçons d'océanographie physique. Ann. Inst. océanogr. Monaco, N. S. *9:* 1–352.

1a. Correns, C. W.
 1937. Die Sedimente des äquatorialen Atlantischen Ozeans. D. Auswertung der Ergebnisse. I. Die Korngrössenvertielung. Wiss. Ergebn. "Meteor," 1925–27, Bd. *3*, Teil 3, Lief. 2: 191–204.

2. Evans, R. D.
1934. Improved technique for determination of radon and radium in liquids. Phys. Rev., (2) *46:* 328.

3. Evans, R. D., A. F. Kip and E. G. Moberg
1938. The radium and radon content of Pacific Ocean water, life, and sediments. Amer. J. Sci., *36:* 241–259.

4. Foyn, Ernst, Berta Karlik, Hans Pettersson and Elisabeth Rona
1939. The radioactivity of seawater. Göteborg. VetenskSamh. Handl., Femte Foljden, (B) *6* (12)*:* 1–44.

5. Hernegger, Friedrich and Berta Karlik
1935. Uranium in sea-water. Göteborg. VetenskSamh. Handl., (B) *4* (12)*:* 1–15.

6. Joly, J.
1908. On the radium-content of deep-sea sediments. Phil. Mag., (6) *16* (XCI): 190–197.

7. Pettersson, Hans
1930. Teneur en radium des dépôts de mer profonde. Résult. Camp. sci. Monaco, Fasc. 81: 1–50.

8. Piggot, C. S.
1933. Radium content of ocean-bottom sediments. Amer. J. Sci., (5) *25:* 229–238

9. 1936. Apparatus to secure core samples from the ocean-bottom. Bull. geol. Soc. Amer., *47:* 675–684.

10. Sanderman, L. A. and C. L. Utterback
1941. Radium content of ocean bottom sediments from the Arctic Ocean, Bering Sea, Alaska Peninsula, and the coasts of southern Alaska and western Canada. J. Mar. Res., *4* (2)*:* 132–141.

11. Thompson, T. G., R. W. Bremner and I. M. Jamieson
1932. Occurrence and determination of iron in sea water. Ind. Engng. Chem., Anal. Ed., *4* (1)*:* 288–290.

12. Urry, W. D.
1941. The radioactive determination of small amounts of uranium. Amer. J. Sci., *239* (3)*:* 191–203.

13. Utterback, C. L. and L. A. Sanderman
1938. Radium content of some inshore bottom samples in the Pacific Northwest. J. Mar. Res., *1* (3)*:* 187–191.

14. Weaver, C. E.
1937. Tertiary stratigraphy of western Washington and northwestern Oregon. Univ. Wash. Publ. Geol., *4:* 1–266.

WATER TABLES IN MARINE BEACHES

By

K. O. EMERY and J. F. FOSTER

University of Southern California
Los Angeles 7, California

INTRODUCTION

Beach slopes, sand movements, and sand composition have been studied by a number of organizations and individuals, but the interstitial water and the possible interrelations between beach characteristics and changing water levels within the beach seem to have received little attention. The present study is an attempt to learn something of those relationships. Surveys were made on three occasions at El Segundo Beach near Los Angeles, and once each at Marine Street Beach and Scripps Institution Beach at La Jolla, California. The authors are aware of little previous related work. Isaacs and Bascom (in press) of the University of California measured some beach water levels in order to learn whether they might serve as indicators of mean sea-level elevation. Zinn (1942) of Rhode Island State College included an examination of beach water levels and water composition in a comprehensive unpublished study of the beach as an environment for micro-organisms. After this report had been written, an abstract by Grant (1946), which outlines much of the thesis of this report, was brought to the authors' attention. In this abstract Grant pointed out that when large waves, heavy rainfall, or small streams inundate a beach back-shore, the water seeps to the surface at the fore-shore, dilating the sediment and moving fine-grained particles. This dilation appears to account for seasonal destruction of beaches and for local degradation of parts of beaches immediately in front of gully mouths.

METHODS

Most of our effort was directed toward measuring the profile of the water table transverse to the shore line and observing the progressive changes of it in response to changing tide level. After an accurate profile of the sand surface was made, a series of pipes 3 to 7 feet long and spaced 5 to 30 feet apart were driven into the sand to a point below the water table. At the crest of the beach the depth to the water table was usually so great that the pipes had to be started from the bottom of an excavation in the beach 5 to 6 feet deep. Each pipe was one inch in inside diameter, tapered and closed at one end, and perforated along the side. The perforations were covered by a fine mesh wire screen which permitted the passage of water but which excluded the sand so that the depth of the water table below the beach surface could be measured with a sounding stick inserted in the pipe. At points nearer the seaward side of the beach the depth to the water table was measured in small holes scooped in the sand. In addition, the position of the edge of the glassy-appearing saturated sand still farther seaward was recorded because it marked the intersection of the water table and the beach surface. Relative tide height and wave height were based on readings on an improvised tide staff driven into the sand. All of the measurements were repeated at half-hour or hour intervals throughout most of a 12-hour tidal cycle.

During some of the surveys, sand samples from various depths were collected and sealed in jars so that water and salt percentage analyses could be made the next day. In the laboratory the samples were weighed, dried at 100°C., weighed again, washed with distilled water, dried, and weighed a third time. The difference in successive weights indicated the amount of water and of salt. At each site a composite sand sample was also collected for mechanical analysis.

A rough measurement of permeability was made by determining the time required to pass a known volume of sea water having a constant head of 30 cm. through a column of sand 15 cm. high and 4.33 cm. in diameter. The sand was obtained by forcing a lucite tube vertically downward into the beach, removing it, and covering the sand-filled end of the tube with a thin cloth.

RESULTS

PROFILES

The profiles measured at all of the sites showed that the water table 20 to 40 feet from the water line lags 1 to 3 hours behind the tide; thus, the water table continues to rise for 1 to 3 hours after the tide begins to ebb and continues to fall for 1 to 3 hours after the tide begins to flood. Apparently this is the result of a slowing of the velocity of the tide wave in passing through the beach sand, which, because of the small interstitial pores, offers much more resistance to passage of the wave than does the open sea. With increasing distance landward the wave lags a longer time and is of progressively smaller amplitude. One would expect, in a beach having high permeability, that the water table amplitude would approximate that of the tide, and that there would be a gradual decrease in amplitude in a landward direction.

Similarly, one would expect that in a beach of low permeability the amplitude of the water table movements would decrease abruptly landward. Examples of the actual changes in the profiles at El Segundo Beach and Marine Street Beach respectively are illustrated in Figs. 1 and 2. These and the other sets of profiles show that during ebb tide the seaward edge of the water table generally slopes seaward, while during flood tide it slopes landward. This is evidently in response to a loss of water from the beach during an ebbing tide and

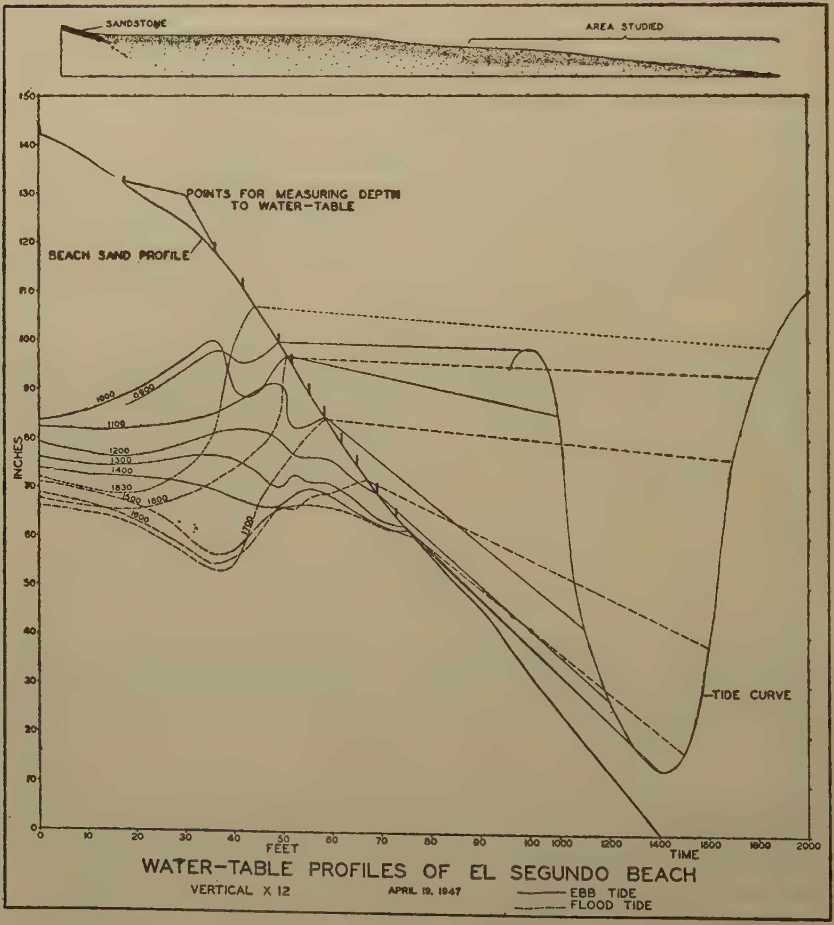


Figure 1. Water table profiles in El Segundo Beach. Insert at top shows the beach with no vertical exaggeration.

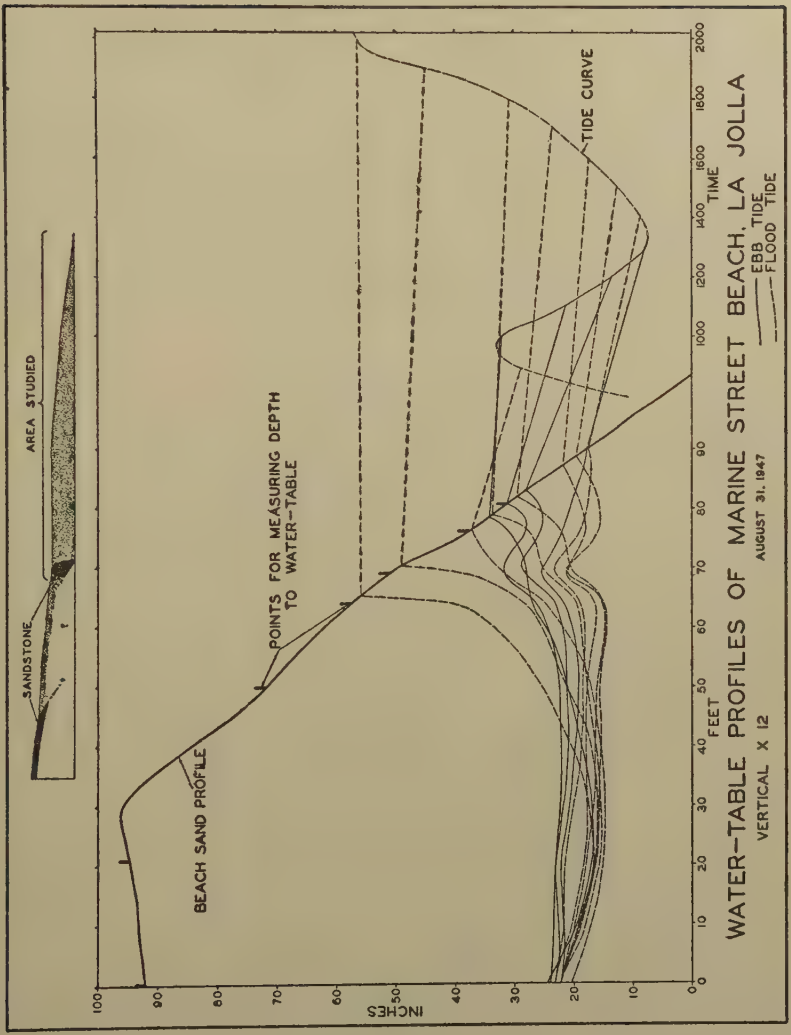


Figure 2. Water table profiles in Marine Street Beach. Insert at top shows the beach with no vertical exaggeration. The sandstone block near the end of the area studied is about 10 feet to one side of the profile.

a gain of water during a flooding tide. The high degree of irregularity in the profiles at El Segundo apparently is produced by local variations in permeability, a conclusion which is supported by the difference of grain size in samples collected at various times (Table I). The profiles at Marine Street Beach are much more regular, and a sand sample from that beach is much better sorted, indicating the probability of higher permeability throughout the beach than at El Segundo. The seaward edge of some of the profiles, taken during an ebbing tide, slope landward (Fig. 2), but this is probably a temporary reversal produced by inflow of water from large waves which washed high on the beach. This wave effect is probably responsible also for the fact that the average seaward slope of the outer part of the profiles at each site taken during ebbing tides is less steep than the average landward slope at flooding tides (Table I). Profiles taken at Scripps Institution Beach were fairly regular, but they showed an amplitude of mostly less than 5 inches in contrast to the 45 to 55-inch amplitudes shown in Figs. 1 and 2. The low amplitude is the result of the small grain size and the low beach slope at Scripps Institution Beach.

PERMEABILITY

Permeability was computed from the field tests, using the usual equation derived from Darcy's Law:

$P = \frac{Q\,l}{A\,h\,T}$, where P is the permeability in darcys, Q is the quantity of water, l is the length of sand column, A is the cross-sectional area of the sand column, h is the hydrostatic head, and T is the time required for the water to pass through the sand. The method is the same as that used for Department of Agriculture soil samples (Fireman, 1944). Resulting permeabilities are given in Table I. Possibly these values are somewhat low because of introduction of air (Christiansen, 1944) during the collection of the sand columns, but air is also trapped in the sand *in situ* (Emery, 1945), decreasing the permeability in nature. In addition, the permeability measurements are based on vertical movement of water through the sand, whereas under actual conditions the water moves both vertically and horizontally. The vertical permeability must be lower than horizontal permeability, because the beaches contain thin alternating coarser and finer layers. Nevertheless, the most permeable sand appears to be that at Marine Street Beach, in accordance with its relative coarseness and excellent sorting; the least permeable sand is that at Scripps Institution Beach where it is fine-grained and less well sorted. It is noteworthy that at least for this set of data, the steeper the beach the greater the permeability, probably because both characteristics are the result of the same environmental factors at the time of deposition. Because Scripps Institution Beach has a very gentle slope, the underlying water table cannot attain the steep seaward slope which its low permeability would permit. The landward slope of the water table at flooding tide also is not steeper than at the other beaches because of the shallowness of the water table and because the tidal effect is partly obscured by water

TABLE I. SOME CHARACTERISTICS OF THE BEACHES AND WATER TABLES

Place	*Date*	*Tide Range*	*Wave Height*	*Beach Slope*	*Slope of Seaward Edge Water Table*		*Median Diameter (mm.)*	*Sorting Coeffic.*	*Perme-ability (Darcy's)*	*Porosity (% volume)*
					Ebb	*Flood*				
El Segundo	4/5	36″	—	4.6°	3.8°	?	0.36	1.18	—	—
El Segundo	4/19	88″	30″	5.9°	3.6°	9.2°	0.38	1.22	—	—
El Segundo	8/16	36″	42″	5.7°	2.8°	7.6°	0.22	1.16	36	26.1
Marine St.	8/31	40″	24″	6.8°	2.2°	14.7°	0.41	1.11	114	26.5
Scripps	9/1	48″	30″	0.4°–1.3°	0.3°–1.2°	1.2° (?)	0.18	1.18	20	27.6

from waves which were able to wash a long distance up the flat beach. This effect of the waves makes it impossible to compute the velocity of flow from the slope of the water table.

SALT CONTENT OF SAND

Seven sets of sand samples were collected from El Segundo Beach and Scripps Institution Beach for water and salt content analyses. These show that the water content of sand drained by a falling water table varies from about zero at the beach surface to saturation below the water table. When saturated, the El Segundo Beach sand contains water equal to 26.1% by bulk volume (Table I). An average water content of about 10% by bulk volume is left behind in the sand as skin or pellicular water by the falling water table. The remaining 16.1% of the bulk volume then is pore space filled by air. Typical proportional volumes of water, air, and solid sand grains at various depths are shown in Fig. 3. Water collected from below the water table at El Segundo contained 3.8% by weight of salts, as determined by evaporation to dryness in the laboratory. Some evaporation occurs in nature, giving rise to higher than normal salt-to-water-ratios and even leading to visible deposits of salt at the surface of the sand. The amount of salt found in the dried sand is a measure of the water content at the time of sampling, plus the amount of water which had evaporated there before the sampling was done. The dashed lines in Fig. 3 indicate the amount of water at various depths, computed from the salt content of the moist sand. At both El Segundo Beach and Scripps Institution Beach the computed water content approximates the measured water content at depth, indicating lack of significant evaporation below about 12 inches in the beach. However, in both areas the salt content proves that considerable evaporation had occurred nearer the surface; in fact, more water evaporated than could have been held originally in the entire pore space of the surface sand. The extra water must have risen by capillarity from some depth in the sand. At El Segundo the top quarter-inch of the sand was cemented into a rigid crust by salt which filled most of the pore space. Such salty crusts are common on sandy beaches.

GEOLOGICAL SIGNIFICANCE

Reference to Fig. 1 shows that during the five hours between the times 1000 and 1500 the water table changed 42 inches at the 40-foot station of El Segundo Beach. It changed only 13 inches at the 0-foot station. We may reasonably assume that a negligible change oc-

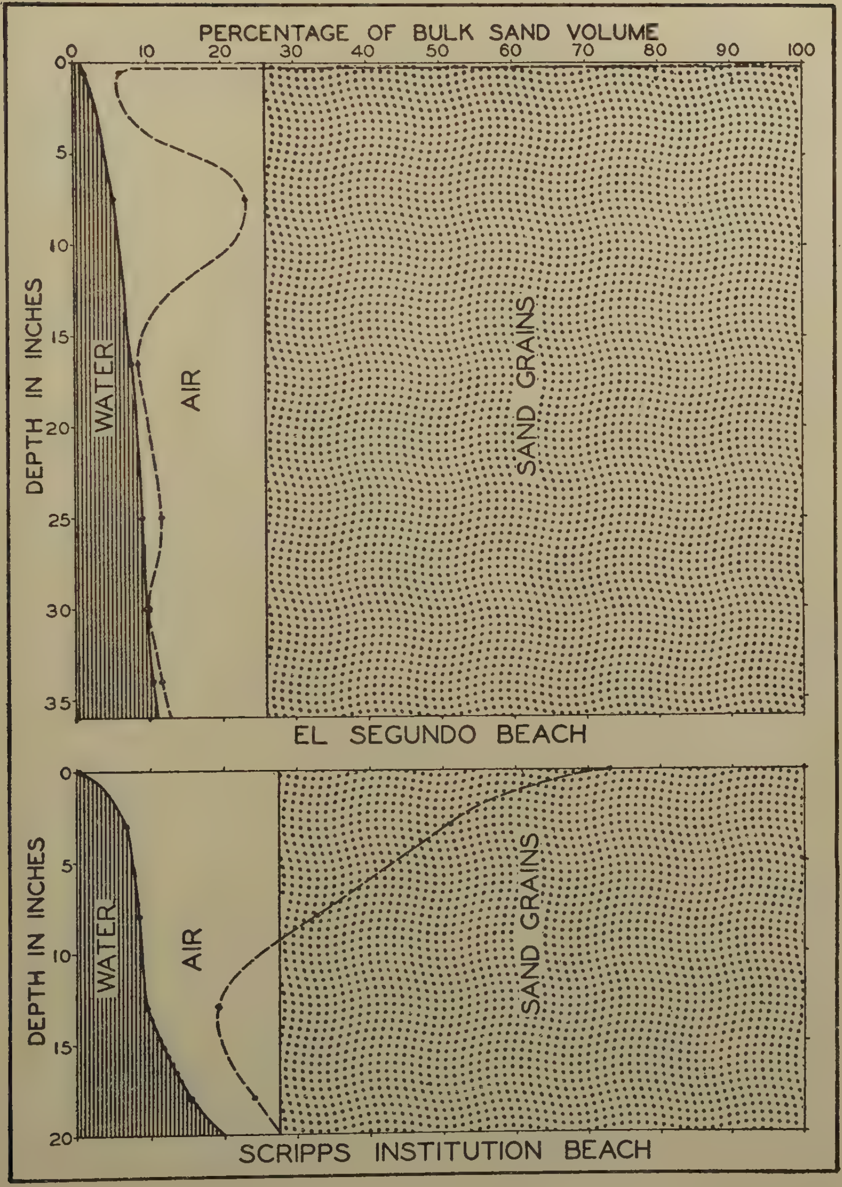


Figure 3. Percentage volumes of interstitial water, interstitial air, and solid sand grains in El Segundo Beach and Scripps Institution Beach. The dashed line is the sum of water still present and water which has evaporated *in situ* as computed from the salt content of the sand.

curred at a point about 60 feet farther landward. Thus, the water table fell through a cross-sectional area of about 200 sq. ft; for a strip of the beach one foot wide, the water table fell through 200 cu. ft. of sand. The sand at El Segundo has a porosity of 26%, which can be assumed to have been fully occupied by water at the high water table position. At the low water table position the sand had a water content of 5.0% by weight, which means that, of the 26% pore space, 10% remained filled with water while 16% then contained air. Thus, water corresponding to 26 minus 10 or 16% of 200 cu. ft. of bulk sand escaped. This amounts to 32 cu. ft., or 6.4 cu. ft. per hour. Only a relatively small percentage of this water escaped upward to the surface by capillarity and evaporated (Fig. 3). It is estimated from Fig. 3 that an average of 0.5 cu. ft. of water per hour rose by capillarity and evaporated at the surface of the transverse strip of sand one foot wide uncovered by the ebbing tide. The remainder, 5.9 cu. ft., had to run down the water table surface and escape through the beach sand. Fig. 1 shows that the escape area during five hours is 25 sq. ft. For any one-hour period the escape area is 5 sq. ft., only 26% of which is open pore space. Accordingly, every hour 5.9 cu. ft. of water passed through 1.3 sq. ft. of area. The escape velocity would have been 4.5 feet per hour, or 0.038 cm. per second. This is approximately the settling velocity of silt of 0.02 mm. diameter. It is obvious that the escape of water from the beach due to tidal change is sufficient to cause the elutriation and eventual loss of silt from the beach, thus contributing to the high degree of sorting of the beach. Mechanical analysis shows the presence of only 0.041% of grains finer than 0.062 mm. diameter (silt) and 0.374% of grains finer than 0.125 mm. diameter (very fine sand). To this tidal effect must be added the escape from the lower part of the beach of additional water which soaks into higher parts of the beach from each wave which runs up beyond the outcropping edge of the water table. The amount of water contributed by waves is unknown and is difficult to measure, but it must be very large during storms, when it certainly exceeds the amount of water from tidal changes. On some beaches, escaping fresh water would add to the velocity (Isaacs and Bascom, in press), although this was not a factor in the present measurements. Consequently, at times, the escape of water from the low part of the beach may be sufficient to exert an appreciable upward force on some of the sand grains composing the beach. Any such tendency toward lifting the sand grains must aid the turbulence of the overlying water to remove the sand and cause some beach erosion. It is noteworthy that Shepard and LaFond (1940) found that during times of spring tides or high storm waves the beach at about mean sea level is eroded and that during neap tides and small waves it is filled again. One of the contributory agents, therefore, may be the water from the falling water table when it escapes through the sand surface.

It is a common observation that at extremely low tides a wide expanse of glassy-surfaced sand may be exposed. Numerous rivulets drain across this surface, beginning at points farther landward than the limit reached by the waves at low tide. These rivulets drain away the water from within the beach which it produced by the lowering water table. Erosion by the running water cuts small meandering channels belonging to the class of beach features known as rill marks.

The rigid crusts formed by evaporation of interstitial water at the sand surface are widespread on marine beaches, but they have not been noted on fresh-water lake beaches examined by the authors. When one walks across a crusted beach the edges of the footprints produced are characterized by a brecciated appearance. Conceivably, such brecciated crusts have been preserved in some ancient beach sands. Their presence might be useful supplementary evidence of "tops and bottoms" in folded rocks, and of marine, or at least salt water, environments of deposition. The temporary hardening of the sand surface by formation of the salt crusts may also be instrumental in the preservation of delicate trails and markings known in ancient sandstones because of its resistance to erosion by waves and wind. Beaches of calcareous debris in tropical regions are commonly cemented into beach limestones. The cause of this cementation is not well established, but it may be partly due to deposition of larger than ordinary amounts of calcium carbonate with the other salts left by evaporation of interstitial water. This cementation forms related but more massive and permanent crusts than those observed on the California beaches.

SUMMARY

Measurements of the changing profiles of the water tables in several beaches throughout a tidal cycle show the presence of an interchange of water between the ocean and the sand pores. Water escapes from the beach during ebb tide with a velocity sufficient to elutriate silt grains, and it may collect into rivulets which erode small rill channels across the beach. Some water is drawn by capillarity to the sand surface where it evaporates, cementing the sand into a rigid crust.

REFERENCES

CHRISTIANSEN, J. E.

1944. Effect of entrapped air upon the permeability of soils. Soil Sci., *58:* 355–365.

EMERY, K. O.

1945. Entrapment of air in beach sand. J. sediment. Petrol., *15* (2): 39–49.

FIREMAN, MILTON

1944. Permeability measurements on disturbed soil samples. Soil Sci., *58:* 337–353.

GRANT, U. S., IV

1946. Effect of ground-water table on beach erosion. Abst. in: Bull. geol. Soc. Amer., *57:* 1252.

ISAACS, J. D. AND W. N. BASCOM

Water-table elevations in some Pacific coast beaches. Amer. Geoph. Union. (In press.)

SHEPARD, F. P. AND E. C. LAFOND

1940. Sand movements along the Scripps Institution Pier. Amer. J. Sci., *238:* 272–285.

ZINN, D. J.

1942. An ecological study of the interstitial microfauna of some marine sandy beaches, with special reference to the Copepoda. Ph.D. Dissertation, Yale University.

INFLUENCE OF THE WATER TABLE ON BEACH AGGRADATION AND DEGRADATION

By

U. S. GRANT
University of California, Los Angeles

INTRODUCTION

Observations for many years of the fluctuation in the width and slope of beaches in southern California have demonstrated that the position of the water table under the beach surface has an important bearing on deposition and on erosion of the foreshore and backshore. During winter storms, in the vicinity of Santa Monica and north along the Malibu Coast, a very rapid destruction of the beaches sometimes results. For example, at Coral Beach (12 miles northwest of Santa Monica) a series of large waves and a high tide may reduce the width of the backshore by as much as 40 feet or more within a few hours and sometimes in less than one hour. The mechanics of such rapid degradation appears intimately connected with the height of the zone of saturation under the beach surface. A high water table accelerates beach erosion, and conversely, a low water table may result in pronounced aggradation of the foreshore. Large waves, with the accompanying excessive turbulence, provide the dynamic factors in rapid beach erosion, but the wetness or dryness of the beach contributes in an important manner to its erosion or aggradation, the explanation of which is the purpose of this paper. A brief synopsis of this principle was presented (Grant, 1946) at the meeting of the Cordilleran Section of the Geological Society of America at Berkeley, April 1946. The recent results of some of Emery and Foster's (1948) observations on beaches add to our knowledge of ground water conditions in the beach environment.

WAVE ACTION ON A DRY FORESHORE

When an incoming oscillatory wave breaks or spills in the plunge zone, propagates itself across the translatory zone and swashes up the foreshore, there is a gradual diminution of velocity. This diminution of velocity on the foreshore, which is a function of momentum, slope angle of the beach, flow depth and the foreshore roughness factor, decreases the transporting power of the uprushing prism of water so that toward the instant of zero velocity most of the sediment transported up the foreshore is deposited on the sloping sand surface. If we have a dry condition of the foreshore (*i. e.*, a water table intersecting the beach surface far below the upper limit of the swash [see Fig. 1A]) water percolates rapidly into the sand above the water table, which reduces the flow depth of the swash and thus reduces the velocity beyond that value due to the operation of the other factors mentioned. The uppermost portion of the swash may all soak into the porous sand, thus depositing all upward transported sediment in that portion of the swash which vanished into the interstices of the beach sand. Furthermore, toward the termination of the time of upward motion of the swash the velocity will become low enough to decrease below the lower critical limit and the flow will change from turbulent to laminar. This change from turbulent to laminar flow can be seen by carefully watching the last moments of the swash on a gently sloping dry beach. Sediment is rapidly deposited when this flow transition occurs, leaving a thin deposit of sediment on the upper portion of the foreshore from

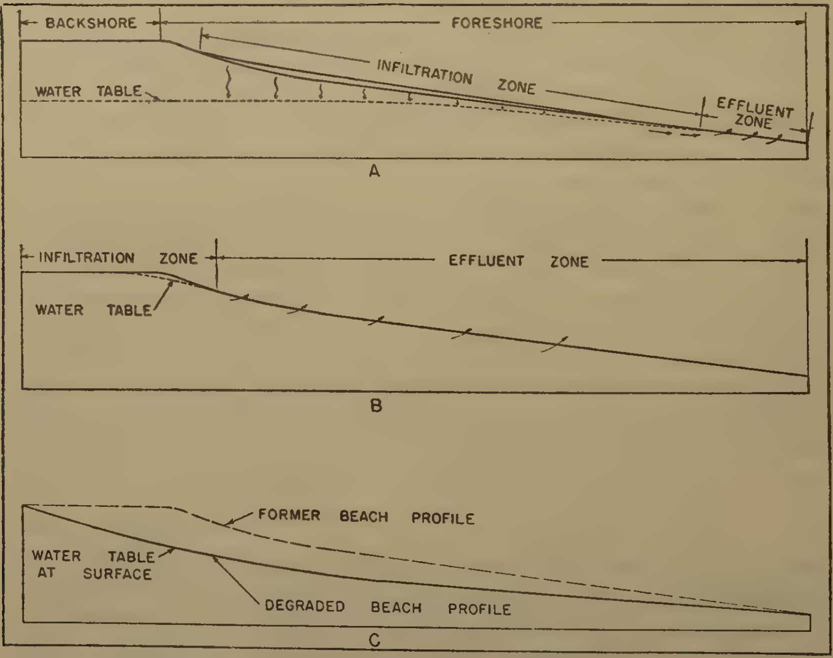

BACKSHORE
FORESHORE
INFILTRATION ZONE
WATER TABLE
EFFLUENT ZONE
A
INFILTRATION ZONE
EFFLUENT ZONE
WATER TABLE
B
FORMER BEACH PROFILE
WATER TABLE AT SURFACE
DEGRADED BEACH PROFILE
C

Figure 1. Hypothetical beach cross sections showing relationship of the water table to the infiltration zone and the effluent zone. A. Low water table with extensive infiltration zone. B. High water table with extensive effluent zone. C. Degraded beach due to water table coincident with surface of beach.

each swash episode. When large waves occur during a high tide the swash may reach beyond the upper limit of the foreshore and flow over the backshore. In this case the water on the backshore may all soak into the dry sand, leaving the sediment load as an accretion on the surface.

After the swash has reached its limit the backwash begins. At first the velocity is low, and laminar flow prevails for a short period. This laminar flow decreases the likelihood of transporting sediment in suspension down the foreshore. The laminar backwash flow persists for a longer time if the slope of the beach is small and if the depth of the film of water is likewise small. Hence, on a gently sloping beach which is dry, the acceleration due to the gravity component (the geodetic or geocentric head) is small and the flow depth, being reduced by percolation into the unsaturated sand, is likewise small. Therefore, under these conditions velocity will be low and laminar flow will persist for a significant time. Unfortunately the author has not yet had an opportunity to determine the Reynolds Number controlling the type of flow of sea water under the conditions of this environment, but the transition may be expressed approximately by the following equation from Jeffreys, 1925 (or Cornish, 1934: 146–147),

$$Q = 300\nu,$$

where Q = the quantity of water crossing unit width in unit time and ν is the kinematic viscosity of water. Since the kinematic viscosity of water is about 0.01 cm^2 per second, this indicates that a film of water one cm. deep cannot flow with a velocity greater than three cm. per second without turbulent flow occurring.

Thus it can readily be seen that a dry beach (low water table) facilitates deposition on the foreshore by reducing backwash flow velocity and thus prolonging the existence of laminar flow. Furthermore, where all of the swash percolates into the backshore or dry foreshore, the sediment transported up the beach by the swash is thus deposited. A gently sloping dry beach is conducive to deposition. As the slope is increased by greater deposition on the upper reaches of the foreshore, the equilibrium slope is gradually reached and backwash velocities prevent further net accretion.

It should be pointed out that, although deposition may occur on the upper foreshore, the lower foreshore, being below the line of intersection of the water table and the foreshore surface, may be subjected to rill erosion, or deposition and erosion may counterbalance each other. This lower saturated foreshore is called here the effluent zone in Fig. 1. A and B.

WAVE ACTION ON A WET FORESHORE

If the water table under a beach is very high and contiguous with the surface of most of the foreshore, a condition favoring beach erosion exists. This situation is illustrated in Fig. 1B. This high water table may be caused by large waves persistently propelling the swash onto the backshore where it has ample opportunity to soak into the permeable sand because surface drainage slopes are absent and surface runoff is negligible. Intermittent wet-weather rills from the coastland back of the beach may bring fresh water to the backshore during heavy rain storms and thus contribute to a high water table.[1] Regardless of source of water, when the beach is in a saturated condition throughout all of the foreshore, the backwash of the waves, instead of being reduced in velocity, is accelerated by addition of water rising to the surface throughout the effluent zone. It can be shown by a simple mathematical demonstration (see Exner, 1931: 378–379) that the mean velocity of flow has the following relationship to slope angle, depth of flow and friction:

$$U = (g \sin i \times h/K)^{\frac{1}{2}}.$$

Here U = mean velocity, g = gravity, i = slope angle, h = depth of flow and K = a friction term. This can be expressed as

$$g \sin i = K\, U^2/h,$$

which shows that if the expression on the left is of constant value, U must increase as h increases, assuming K is constant or varies more slowly than the other variables. Hence, the increased volume of backwash has a high velocity, and turbulent flow exists. This enhances erosion of the foreshore. The addition (instead of subtraction) of water to the backwash by ground water escaping to the surface of the foreshore not only increases backwash flow velocity but dilates the sand and propells the finer grains up into the turbulent flow. Emery and Foster (1948) consider this factor.[2]

Another feature is sometimes produced by the increases in volume and velocity of the backflow of water on the foreshore. The backwash is often met by an oncoming translatory wave approaching the lower foreshore. If the backwash has a velocity greater than $(g \times h)^{\frac{1}{2}}$, the approaching water cannot progress against the current and a hydraulic jump or a roll wave with high turbulence occurs. This excessive turbulence not only prevents deposition but removes all but coarse particles whose settling rates in sea water are so great that they are not readily transported. If this hydraulic jump or roll wave is often repeated in the same position, a small scour channel lined with coarse sand or gravel persists at the lower edge of the foreshore parallel to the trend of the beach. The scour channel due to the hydraulic jump appears to be developed only at the base of a steeply sloping foreshore.

---

[1] Direct rainfall on the backshore is apt to contribute only a little in elevating the water table unless an unusually thin sand prism rests upon an impermeable substratum and unless the rainfall is very heavy.

[2] The escape of ground water on the foreshore may reduce the value of the friction term, but this influence is difficult to evaluate. In any event it would serve to increase the erosion effect.

The aggrading or degrading of a beach, and the value of the angle of beach slope, is a function of several variables, one of which is the position of the ground water table under the beach. A degraded beach, shown in Fig. 1C, may be in equilibrium during a saturated condition but may be aggraded when a dry condition and normal wave action returns.

REMOVAL OF BEACH BARS

Along the southern California Coast there are a number of aggraded valleys occupied by intermittent streams whose mouths are completely blocked from the sea by a beach bar during the dry season. When the wet season comes and stream flow begins, the runoff water is impounded back of the beach. If stream flow begins with a large volume of runoff, the water in the lagoon back of the beach may rise rapidly above the crest of the sandy beach bar and cut a channel to the sea. If the water in the lagoon rises slowly, then the beach water table at the blocked river mouth may rise enough to conspire with the first episode of large waves to greatly accelerate beach erosion locally and thus facilitate the removal of the beach bar. Quantitative studies of this phenomenon have not yet been made, but local narrowing of beaches opposite these bars have been seen where water is impounded in the valley mouth lagoons.

EFFECT OF THE WATER TABLE ON ARTIFICIAL SAND SPREADING

The dilating effect of escaping ground water in the effluent zone of the foreshore, the enhanced erosion there due to this decrease in compaction of the sand grains, and the increased velocity of the backwash on the foreshore by increase in flow depth, have an important bearing on the method of spreading sand on a beach by means of sluicing sand from sand dunes or from a dredger. On a narrow beach with a sensitive water table the sluice water carrying the sand must be effectively spread over the beach to prevent erosion counteracting the effect of the artificially added sand. Failure to consider these factors contributed greatly to the negligible aggrading effect of pumping sand from the aggraded beach back of the Santa Monica Yacht Harbor (Grant and Shepard, 1939: 803–805) to the narrow degraded beach a few hundred yards south of that structure during 1939.

REFERENCES

Cornish, Vaughan
1934. Ocean waves and kindred geophysical phenomena. University Press, Cambridge. 164 pp.

Emery, K. O. and J. F. Foster
1948. Water tables in marine beaches. J. Mar. Res., *7* (3): 644-654.

Exner, F. M.
1931. Zur Dynamik der Bewegungsformen auf der Erdoberfläche. (Gerlands) Beitr. Geophys., Suppl. Bd. *1:* 373–445.

Grant, U. S.
1946. Effect of ground-water table on beach erosion. Bull. geol. Soc. Amer., *57:* 1252.

Grant, U. S. and F. P. Shepard
1940. Shallow-water sediment-shifting processes along the southern California coast. Proc. 6th Pacif. Sci. Congr. (1939): 801–805.

Jeffreys, Harold
1925. The flow of water in an inclined channel of rectangular section. Philos. Mag., *49:* 793–807.

EVIDENCE OF WORLD-WIDE SUBMERGENCE[1]

By

FRANCIS P. SHEPARD

Scripps Institution of Oceanography
University of California
La Jolla, California

INTRODUCTION

At the beginning of the century geologists were discussing the idea that the glacial period was the result of enormous vertical movements of the continents. Uplifts were said to have cooled the climates, bringing on glaciation, and subsequent submergence caused the glaciers to melt. This hypothesis was supported by the finding of deep canyons on the sea floor which indicated the former elevated condition of the continents. The idea was later abandoned, partly because of the very limited amount of information concerning submarine canyons but largely because the glacial period was found to have a series of ice stages or epochs with intervening times of relatively warm climate. The alternate elevation and depression of the continents to a degree sufficient to produce widespread glaciation, all during about one million years, appeared to be completely out of accord with the history which geologists had been building from their accumulating evidence.

The abandonment of the diastrophic hypothesis seems to have resulted in shelving the subject of submarine canyons, so that for a quarter of a century there was almost no mention of them in the literature. This neglect terminated when the amazing advances of hydrographic surveys of the past two decades revealed the presence of abundant submarine canyons of huge dimensions off many coasts of the world. While many geologists have considered these canyons as having been cut by running water, and therefore being indicative of great changes of level, others have developed hypotheses by which they attempted to explain the canyons as having been formed under the ocean. These explanations have proved satisfactory to many geologists, but as oceanographic research has progressed it has become increasingly evident that the submarine canyons are remarkably like the river-cut canyons of the land. During the same time, an abundance of field observations has supplied independent evidence that there has been a very general submergence of continental borders, oceanic islands, and perhaps even of the deep sea bottom.

---

[1] Contributions from the Scripps Institution of Oceanography, New Series, No. 387.

SUBMARINE CANYONS

Nature of Canyons. Because of the implications of enormous changes of level resulting from the world-wide distribution of the canyons, if they prove to be submerged stream valleys, some assurance should be provided that they are truly stream-eroded rather than the product of processes which can operate on the sea floor. If one considers the number of illustrious scientists who have offered hypotheses attempting to prove a submarine origin, skepticism of the stream origin would be natural. However, if one studies the detailed charts of the canyons and examines the evidence which has come from canyon exploration by dredging, coring, bottom photography, current measurements, and diving, the skepticism should not be as great. In fact, a study of the field data leaves one with the impression that the various authors of hypotheses of submarine origin were describing somewhat imaginary undersea valleys.

Submarine canyons have been described from many parts of the world (Fig. 1). Canyons have been discovered on virtually all well-explored continental slopes except where there are indications of relatively recent faulting or where the slopes are very gentle. Further exceptions may exist off coasts which now have very slight precipitation.

Some writers have been inclined to consider that the canyons were entirely different in character in different areas. For example, they have been quite willing to concede that the California canyons were the result of stream erosion, while they considered that the canyons in most other areas were of a different nature. In the present article the canyons can not be discussed adequately, so that reference should be made to a recent comparison of canyons from all parts of the world (Shepard, 1948: 210–229). So far as evidence is now available, the conclusion from this work is that the canyons are essentially the same the world over. They cut the submarine slopes as deep notches with *V*-shaped transverse profiles. They have narrow floors sloping almost continuously away from the lands. They have twisting courses and tributaries entering with the dendritic pattern characteristic of land canyons. The walls are precipitous and have proved to contain rock exposures in almost all of the areas from which samples are available. Many of the rocks are of a type resistant to erosion. The majority of the canyons extend shoreward across most of the continental shelf and many of them enter estuaries. Others, however, have their heads on

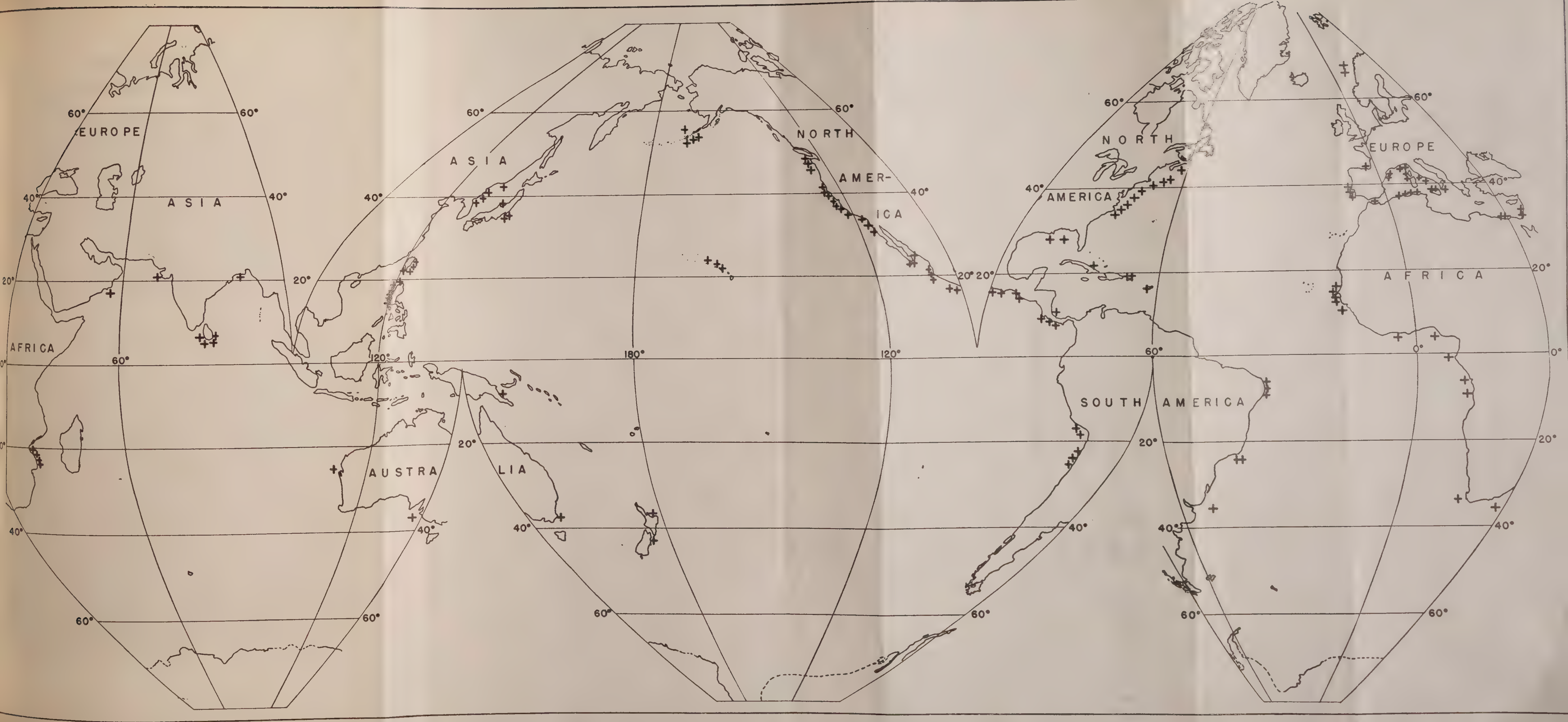

EUROPE
ASIA
AFRICA
ASIA
NORTH AMERICA
AUSTRALIA
NORTH AMERICA
SOUTH AMERICA
EUROPE
AFRICA
60°
40°
20°
0°
20°
40°
60°
60°
120°
180°
120°
60°
0°

Figure 1. The world distribution of submarine canyons. Each cross indicates one or a group of canyons. Many areas lack adquate soundings for the determination of canyons.

the outer shelves far from land or on the upper portions of the continental slopes.

The canyons can be traced outward to depths of many hundreds of fathoms. The outer terminations of most canyons can not be ascertained from available evidence. In some cases they show a change from a *V* shape to a trough shape at depths of the order of 1,000 fathoms. Off southern California some of them appear to end with flats or submerged deltas at depths which are approximately those of the rim of the basin into which the canyon has been cut.

The most completely studied canyons have proved to be identical in character to the land canyons which are directly inside (Fig. 2). Frank Haymaker of the Scripps Institution has dived into and explored the head of two canyons at La Jolla, California. He found typical stream erosion features, such as staircase canyons with waterfall drops, hanging valleys, tributaries entering at grade, vertical and even overhanging rock walls, and cobble covered floors. The canyons on land inside have all of these same features.

Possible Submarine Origins. So many types of submarine origin have been suggested for canyons that it is difficult to discuss each adequately in the present paper. A few of the principal hypotheses will be mentioned briefly.

The idea that the canyons are the result of faulting or rifting (Wegener, 1922: 177, 178) seems to have been developed without any apparent knowledge of their character. Fault valleys are obviously unlike the canyons just described. The former are in general straight-sided, flat-floored, free of tributaries, and have deep basin depressions. The faulting origin is not now seriously considered by any one acquainted with the subject of submarine canyons.

The idea of submarine artesian spring sapping (Johnson, 1939) seems also to be based on very little other than the conviction that other hypotheses are inadequate. The idea depends largely on the presence of good aquifers outcropping on the continental slopes, with the circulation of water undermining the slope and thus forming gullies which might be enlarged by landslides. However, the canyons are by no means confined to the slopes, since most of them, as explained previously, cross the shelves and many extend into estuaries. Furthermore, the rock on the walls of the canyons does not favor the aquifer idea. Nor do the canyon floors have the large sinks or basins which would be developed by underground water. The possibility for artesian circulation under many of the areas where canyons exist is negligible. For example, in some places faults intervene at the coast preventing any appreciable circulation of this sort.

The conception of origin by tsunamis (tidal waves) (Bucher, 1940) seems to be equally unfortunate. The great waves, set up rarely by submarine faulting, are virtually confined in their effects to the shore and coast. Although the tsunamis develop strong currents in shallow water along the shore, their effectiveness decreases rapidly with depth. In long waves of the Stokes type, the horizontal bottom velocity under the crests of the waves where it reaches a maximum of short duration, is given by the formaula: $Q = CH/(2h)$, where Q is the velocity of the bottom current, C is the wave velocity, H is the wave height and h is the depth of water. To obtain the wave velocity, use is made of the formula $C = \sqrt{gh}$, where g is the acceleration of gravity. Using these formulae and considering a very large tsunami, we find that the maximum currents would have values approximately as follows:

Fathoms	*Cm/sec*
5	95
10	62
100	11
300	4.5
500	3

These figures indicate how powerless tidal waves are in producing erosion in deep areas. Furthermore, several areas which have an abundance of tsunamis have no canyons, as for example the Sanriku Coast of Japan.

The only hypothesis which has been supported in print by more than one geologist is that of cutting by density currents (Daly, 1936). According to this idea heavy muddy water moves down slopes, thereby excavating the canyons. The hypothesis has some appealing features, since these density currents are actually known to exist, at least in lakes. Model studies show how the currents produced by heavily weighted liquids could cut gullies into soft materials. Beyond this point the hypothesis has been built on highly hypothetical claims. It has been said that density currents flow along the bottom of Lake Mead at the rate of four miles an hour. Actually, the basis for this rate has been shown by the U. S. Reclamation Service to be groundless, and the present studies of Lake Mead by the U. S. Geological Survey show that the density currents are so feeble that they are allowing the rapid deposition of clay at the front of the delta of the Colorado River where the currents are best developed. In fact, at the places where density currents should be at their best, deposition is going on rapidly. This being the case, one should hesitate to believe that canyons thousands of feet deep cut into hard rocks could be due to such currents.

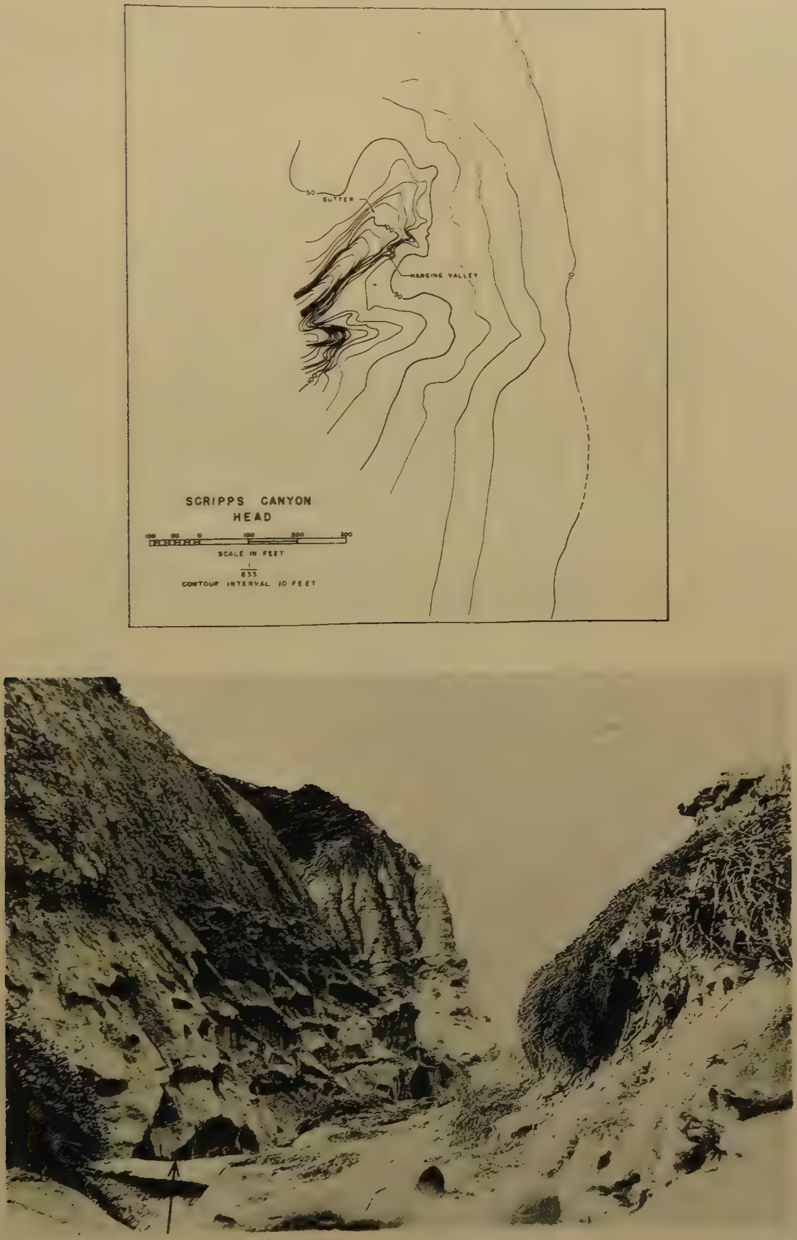

50
GUTTER
100
HANGING VALLEY
50
10
100
SCRIPPS CANYON HEAD
100 50 0 100 200 300
SCALE IN FEET
1/833
CONTOUR INTERVAL 10 FEET

Figure 2. Comparison between the submarine canyon head (upper) and an adjacent land canyon (lower) at La Jolla, California. Both canyons have hanging valleys and precipitous walls. The arrow in lower illustration indicates the edge of a 40 foot vertical cliff.

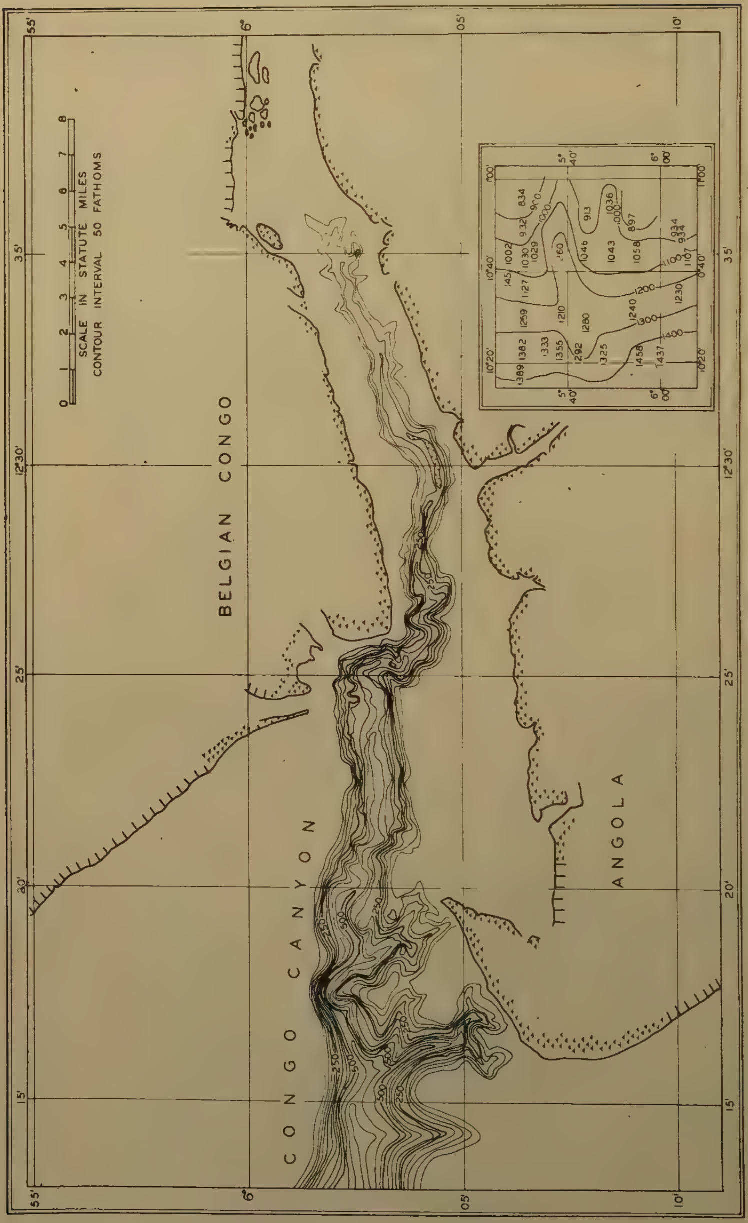

0 1 2 3 4 5 6 7 8
SCALE IN STATUTE MILES
CONTOUR INTERVAL 50 FATHOMS
BELGIAN CONGO
ANGOLA
CONGO CANYON
55'
6°
05'
10'
35'
12°30'
25'
20'
15'
250
500
10°40'
10°20'
5° 40'
6° 00'
1400
1300
1200
1100
1000
900

Figure 3. The deep estuary at the head of Congo Canyon. The inset shows the outer canyon terminus.

Stream Origin and Landslide Clearing. Since none of the existing marine origin hypotheses can explain the characteristics of the canyons and since stream-cut canyons are essentially identical in character with the submarine types, the subaerial origin appears to be by far the most promising. It is possible, however, that currents now completely unknown have been the cause. Certainly, the currents now existing on the canyon floors are far too weak to be held responsible. These currents have been determined by current meter observations, and their nature is confirmed by diver reports. They are barely strong enough to keep the finer sediment in suspension except under unusual circumstances when they cease entirely.

Deposition is in progress in the canyons, as is now well known from the California studies. The canyon heads would soon be filled by this deposition if it were not for the instability of the fill. The mud in the sediments and the organic debris, particularly kelp and eel grass, apparently favor slumping along the inclines at the canyon heads so that the fill slides seaward periodically. Evidence both of the fill and of the slides comes from the repeated soundings along carefully surveyed lines. Slides have also been detected by sudden deepening at the ends of piers which had been built into the canyon heads. To date, there is no evidence that these slides do more than remove sediment from the canyon floors. It is possible that some enlargements may occur, particularly in the nature of widening, but it seems almost impossible that the huge canyons could have been cut thousands of feet into the solid rocks underlying the continental slopes by landsliding alone. The extension of a canyon 50 miles across the continental shelf and up into an estuary (Fig. 3) by landsliding is obviously out of the question. The shapes of the canyons are by no means those of landslide scars. The latter are for the most part cirque-like and have greater width than length. Also, landslide scars have irregular floors with many basins along their courses. The sliding associated with the canyons must be preservative and not formative.

Implications of the Stream Origin Hypothesis. Almost all objection to the subaerial origin of the canyons has come from concern over the implications which such an origin would have. If the canyons were found only off certain unstable coasts such as California or Japan, it might be easy to convince geologists that the canyons were submerged by regional earth movements, but the finding of these canyons off supposedly stable coasts as well as unstable coasts without indication of any preference for the one or the other has caused considerable uncertainty. It is not only necessary to account for the submergence of the continental margins by a mile or more, but also for the deep drowning of insular canyons. We find, for example, that the canyons are well developed in the mid-Pacific Hawaiian Islands (Fig. 4). Here they are found off the deeply eroded windward coasts of the older islands, but so far as is known they are not present off the coasts where there have been recent lava flows.

It might be argued that the canyons have been submerged from time to time in the remote geological past, but most of the evidence which we now possess favors relatively recent submergence. The existence of narrow gorges with vertical and overhanging walls is not indicative of very great age, since the burrowing and boring of organisms on the ocean bottom should be capable of reducing these slopes to moderate inclinations over the course of millions of years. Furthermore, some canyon walls are definitely known to contain Tertiary formations, ranging from Eocene to Pliocene. This suggests that at least the inner parts of many of the canyons were cut during the Pleistocene or Recent. The Tertiary formations incidentally are found off both the east and west coasts of the United States, so that the implications of recency are not confined to unstable coasts. A valley has been found to extend to a depth of 350 feet below sea level in the delta of the Mississippi. This valley was traced to a depth of 1,000 feet at the coast by an oil company using geophysical methods of exploration and later was traced across the shelf nearly to the point where the head of Mississippi submarine canyon appears on the charts. The age of the excavation of the filled head within the delta has been definitely determined by Fisk (1944: 11) to be late Pleistocene. By implication, the canyons which are cut into the edge of Georges Bank off the New England Coast are contemprary with the stand of the glaciers on these banks, since otherwise it would be difficult to imagine sufficient drainage from the small water sheds of the banks to cut canyons that are comparable in size with the canyons off the large river valleys from the Hudson to Chesapeake Bay. On the other hand, there is no evidence at present which would prove that more than the inner portions of the canyons down to depths of about 1,000 feet were cut during this late geological epoch. The deep outer portions may have been excavated at various times during the past and their filled heads reopened by excavation during the glacial period.

Conclusions Concerning Submarine Canyons. At the present stage of our canyon investigations we have found substantial evidence that the canyons are river-eroded whereas we have so far failed to find indications of submarine origin. The canyons are apparently present on both insular and continental slopes except where recent faulting or vulcanism has interfered, or where the slopes are so gentle that valleys

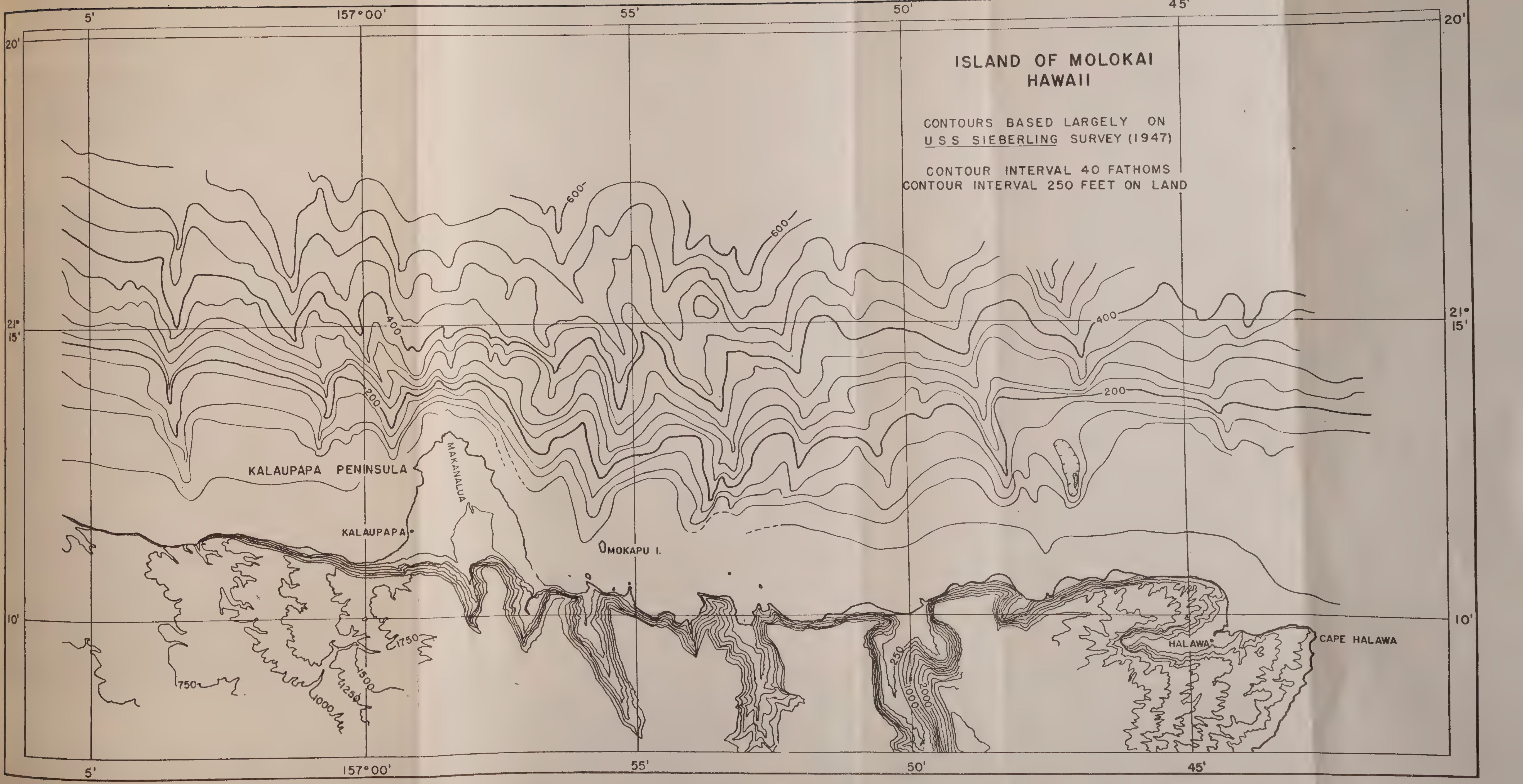

ISLAND OF MOLOKAI
HAWAII
CONTOURS BASED LARGELY ON
USS SIEBERLING SURVEY (1947)
CONTOUR INTERVAL 40 FATHOMS
CONTOUR INTERVAL 250 FEET ON LAND
5'
157°00'
55'
50'
45'
20'
21° 15'
10'
600
600
400
400
200
200
KALAUPAPA PENINSULA
MAKANALUA
KALAUPAPA
OMOKAPU I.
HALAWA
CAPE HALAWA
1750
1500
1250
1000
750
250
1000
2000

Figure 4. The submarine canyons north of Molokai, Hawaiian Islands. Note relation of sea and land canyons.

would have insufficient gradients for landslides to dispose of the fill. This universality of canyons suggested the possibility that the cause of the canyon cutting and submergence was a change in sea level rather than a world-wide movement of the crust. If this change has taken place, it should be registered in other ways. The rest of this paper will deal accordingly with other lines of evidence.

CORAL REEFS

The suggestion of wide scale submergence came from coral reef studies at an even earlier date than from the submarine canyons. As is well known, reef corals grow only at shallow depths. Darwin (1842) first proposed that the atolls represented the growth of coral masses on top of submerging volcanoes. This idea has been as hotly disputed as has the stream origin of canyons. Foremost among the contending hypotheses has been that of Daly (1910). He suggested that wave planation of island tops occurred during low stands of the sea accompanying high latitude glacial stages. He attributed the atolls to the growth of coral crescents around the rims of the levelled platforms during the rising sea level accompanying glacial melting.

Fortunately, it has been possible to apply fairly conclusive tests to these two contending hypotheses. If the wave planation idea is correct, one would expect that boring into the atolls would expose the wave abraded platform at moderate depths. On the other hand, the Darwin hypothesis would be supported by finding shallow water formations extending to great depths. We are now in possession of evidence from four borings into atolls at scattered points in the western Pacific. All of them show the same result, namely, that no wave abraded platform was discovered. The recent boring on Bikini (Ladd, Tracey, and Lill, 1948) provides the best evidence. It extended to 2,500 feet and shows, so far as has been found, nothing but shallow water formations for the entire depth. The core is in no way indicative of being talus material later grown over by an advancing reef, as Daly has contended. The boring partially confirms the geophysical evidence which suggests that the igneous core is from 6,000 to 13,000 feet below the surface (Dobrin, *et al.*. 1946). Another relatively recent boring on an atoll northeast of Borneo (Kuenen, 1947) revealed deposits of coral limestone and lagoonal marl down to a depth of 429 meters, the bottom of the hole. Relative to a somewhat earlier drilling on the former Japanese island of North Borodino, a slightly elevated atoll east of Okinawa, the Japanese (Hanzawa, 1940) reported 431 meters penetration through calcareous sand and limestone with shallow water foraminifera along the entire length. The earliest and best known of the cores was from the atoll of Funafuti. According to what is known of this core, the entire 1,114 feet is indicative of shallow water origin and of materials like those of the present reef or of the lagoon. The investigation of the core was made so long ago that the results are now questioned. Some of the core is to be restudied in the near future to bring the data into line with what is now known about ecology of organisms.

Indications of submergence are less pronounced in the case of the borings of the Queensland Reef off the Australian coast (Richards and Hill, 1942). Here the reef proves to have a thickness of only about 400 feet. It does not overlie a wave-abraded platform but rests on a coarse grit formation suggestive of shallow water deposition. At Bermuda the borings revealed weathered rock at 245 feet, sand and gravel at greater depths, and finally volcanic material at 560 feet, but no definite wave-cut platform. The results of the deep borings in the Bahamas by oil companies are not yet published, but it is known that they have penetrated some 14,000 feet of limestone, all apparently of shallow water origin.

According to Kuenen (1933), evidence of submergence of the coral islands comes from the steep slopes which surround some of them. These slopes are considered too steep for talus and are interpreted as due to upgrowth as the reefs submerged. Also, Kuenen has called attention to the existence of atolls rising from deep platforms which have no evidence of vulcanism, so that the atolls have evidently grown up as the platforms submerged. Davis (1928) introduced various other lines of evidence, too numerous to list, suggesting submergence of the coral islands. Perhaps the most impressive of these is the wide distribution of atolls in the central Pacific in areas where remnants of volcanic islands are not found (Fig. 5). It would be most surprising if volcanic islands, many of which must have been as large as the Hawaiian group, have been so completely bevelled by the waves as to leave nothing but platforms. Obviously, this abrasion could not be accomplished during short glacial epochs. Furthermore, if reefs have simply grown on the tops of shallow banks, it is difficult to account for the absence of an equal number of shallow reefless banks in the zones north and south of the coral seas.

Age of Submergence. The time of submergence of the coral islands may not be the same as that of the submarine canyons. The Bikini boring has shown definitely that there are Tertiary formations at depths of 900 feet or less. Pliocene is underlain in turn by Miocene. Similarly, the deep Bahama boring passes down through most of the Tertiary. These borings suggest a gradual submergence and by them-

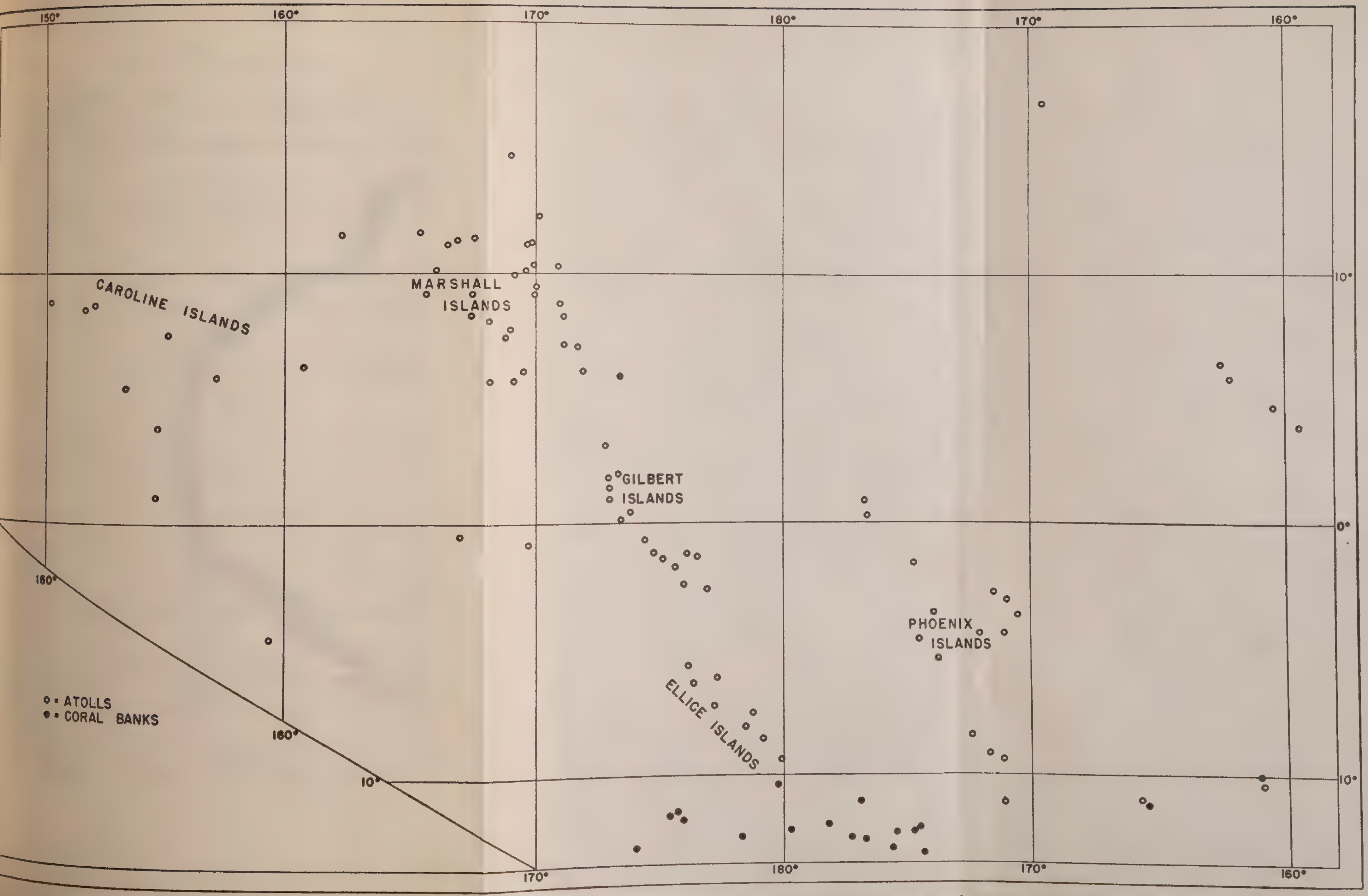

150°
160°
170°
180°
170°
160°
10°
0°
10°
CAROLINE ISLANDS
MARSHALL ISLANDS
GILBERT ISLANDS
PHOENIX ISLANDS
ELLICE ISLANDS
180°
160°
10°
170°
180°
170°
160°
○ = ATOLLS
● = CORAL BANKS

Figure 5. The distribution of atolls and coral banks in the Southwest Pacific. All of these islands lack volcanic rocks so far as is known, and except in the Caroline Islands volcanic islands are missing from these groups.

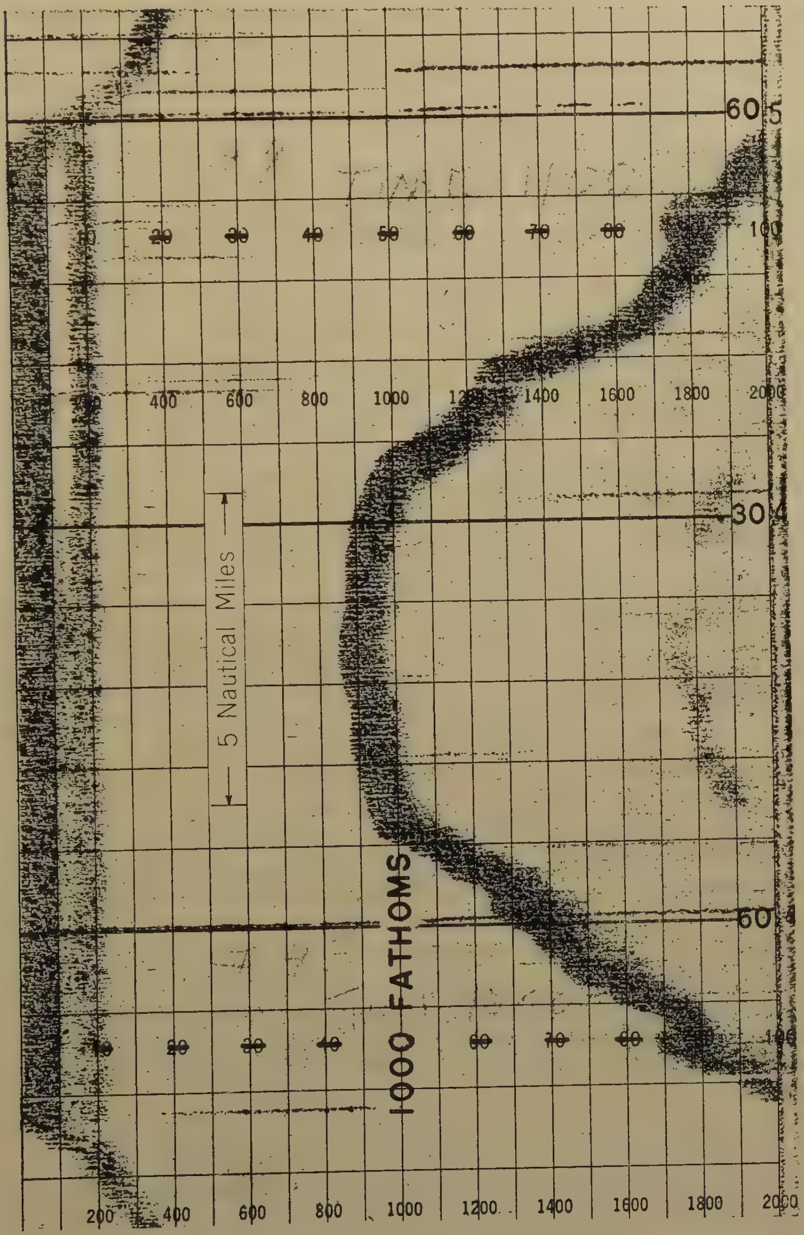


Figure 6. A fathogram of a seamount with terraces in the Gulf of Alaska. Center of seamount is at Lat. 52° 34′ N, Long. 151° 19′ W.

selves are not indicative of sea level changes. The deep Funafuti boring was apparently all in Pleistocene and may be contemporaneous with and related to the Pleistocene portions of the canyons.

FLAT-TOPPED SEAMOUNTS (GUYOTS)

An interesting discovery came from the use of the recording echo sounding devices by the Navy during World War II. A series of flat-topped, deeply submerged mountains was found in the West Central Pacific. We are indebted to Hess (1946) for most of the available information on these bevelled mountains, which he has termed "guyots." The flat tops were found to vary considerably in level, but about 50 per cent of those in the Southwest Pacific lie close to 800 fathoms in depth. On the other hand, the seamounts in the Gulf of Alaska and off the California Coast have summit flats at variable depths averaging around 400 fathoms.

The seamounts off California have terraces. The same appears to be true of the submerged mountains in the Gulf of Alaska (Fig. 6). Both the terraces and the flat tops are difficult to explain except as wave-bevelled platforms, or in the West Central Pacific as submerged coral banks. In either case the surfaces must have been brought close to sea level so that we appear to have further evidence of large submergence. Some flat-topped mountains were discovered in recent work by Maurice Ewing in the Atlantic. It should be pointed out, however, that in the Antarctic R. S. Dietz obtained fathograms during the 1946–1947 expedition and failed to find any flat tops to the mountains or ridges which he discovered. Probably information will become available soon from many other areas.

Information concerning the materials of the seamounts was obtained in a recent cruise of the Navy Electronics Laboratory vessel EPCE(R) 857. During this cruise Twin and Erben Banks, 580 and 860 miles off southern California, were explored under the direction of R. S. Dietz. Dredgings procured volcanic rocks with rounded cobbles from each bank at depths of about 350 fathoms. On Erben Bank the cobbles were embedded in a crust of manganese which had a thickness of 4 cm. If Pettersson's (1945) estimate of the rate of radial growth of manganese nodules of 1 mm/1,000 years is correct, the suggestion is that Erben Bank has been submerged about 40,000 years. This, of course, assumes that manganese accretion began as soon as the bank became deeply submerged, which may not have been the case. Indirect evidence of the age of guyots in the Marshall Island area comes from the boring at Bikini. The atoll appears to be built on one edge of a guyot platform, so that the flat top of the seamount could be considered as older than the formations which cover it. Since the 2,500 foot drilling has penetrated to Lower Miocene, the guyot surface is apparently pre-Miocene. Hess (1946) suggested that the guyots of the Southwest Pacific were bevelled in the pre-Cambrian and that they have been sinking slowly ever since due to oceanic sedimentation which has weighed down the floor of the ocean. This is an interesting suggestion, but seems rather unlikely in view of other evidence of what must be much more recent submergence.

COARSE SEDIMENTS AND SHALLOW WATER FORMS OBTAINED FROM GREAT DEPTHS

Geologists have generally assumed that coarse detrital sediments could be deposited only in shallow water, although it was recognized that some coarse material could be transported to a deep water environment by rafting or even by landsliding. Rafted sediment dropped into deep water can usually be distinguished by the finding of coarse debris mixed with typical deep-water deposits. Landslide deposits are largely confined to the base of relatively steep slopes, but they are otherwise not easy to distinguish in bottom samples. In general, one might expect to find coarse debris mixed with sediments in the vicinity of great glaciers discharging into the oceans, as, for example, around Antarctica and Greenland. Such finds have been numerous and the deposits are called "glacial marine." However, the finding of coarse sediment in deep water at considerable distance from the ice fronts should be rare because of the limited amount of transporting facilities other than icebergs. The debris drifted by vegetation mats, tree trunks, and kelp must constitute an insignificant proportion of the deep sea sediment which is contributed by marine life and colloidal material carried from land by currents. Therefore, it is interesting to find that coarse sediments are not uncommon among the few ocean cores now available. Sand layers have been obtained in cores from the deep Atlantic far from any shoal areas, according to H. C. Stetson and Maurice Ewing. Cross-bedded sand was found in the Gulf of Mexico at over 2,000 fathoms according to Trask, Phleger and Stetson (1947). According to R. Revelle, cores obtained by Scripps Institution in the deep basins off California and Lower California have also shown sand layers. Gravel was found in the Dutch East Indies at 20 out of 385 stations, in depths ranging from 192 to 1,400 fathoms (Neeb, 1943: table 2 on pp. 72–85, also 225–227). It is possible that this coarse material was deposited at great depth, but the alternative of submergence must be considered.

Other evidence comes from the rounded nature of the rocks dredged from ridges and seamounts. Since the violent turbulence of shallow water conditions is necessary to produce rounding, rocks formed on the deep ocean floor by vulcanism would rarely, if ever, be rounded. Some of the rounded rock obtained in the North Atlantic may have been introduced by ice, but the cobbles on banks and seamounts off California (Shepard, 1941: 1869, pl. I) are in all probability an indication of submergence. It is unfortunate that so few of the marine highs have been sampled, but it is probably significant that a large per cent of those which have been explored have yielded rounded rocks. If the percentage holds good for future exploration, this source of information will add substantial evidence to the other lines. Possible confirmation of the rounded rocks in dredgings comes from the numerous chart notations of gravel in deep water.

The coring operations under the direction of Pettersson (1946) in the Mediterranean yielded evidence that the deep parts of that sea may have been shallow at not too remote a period. The cores contained layers with abundant molluscs, the latter being almost entirely confined to a shallow water habitat. Dangeard (1928) reported the discovery of lamellibranch shells with gastropod borings in 660 fathoms west of the Strait of Gibraltar. Beebe (1932: 39) reported a dredging of shallow water shells along with rounded gravel at 1,200 fathoms on a ridge off Bermuda. Kuenen (1933) referred to the dredging of recent coral reef material in 550 fathoms in the Ceram Sea. Of possible less significance was the discovery of littoral shells at depths of 1,420 fathoms off the Norwegian Coast (Hull, 1912). Much more exploration is required before the faunal evidence can be properly evaluated.

INSULAR FAUNAS AND FLORAS

The submergences indicated by the preceding lines of evidence should be reflected also in the faunas and floras of islands, provided that the submergences had been of sufficient extent to destroy former land bridges to the mainland or to adjacent islands. The literature relative to island connections is large and some of it contradictory. Some extremists claim former connections of many of the true oceanic islands with the mainland, but the general concensus of opinion seems to be that islands separated from the mainland by deep and wide channels, that is, channels over 1,000 fathoms deep and 100 miles or more in width, have waif faunas and floras, either rafted to or blown onto the islands. On the other hand, the life forms of islands on the continental shelves are for the most part clearly related to the adjacent continents and require land bridges.

Between the extremes of true oceanic islands and continental shelf islands are the cases where there are channels of depths between 100 and 1,000 fathoms. Among the best known of this group are the islands of the "continental borderland"[2] off southern California. Here the sills of the channels range from 150 to 450 fathoms. The faunas of all of the larger of these islands include endemic species which are difficult to explain without land bridges. The presence of endemic foxes is particularly impressive, since it seems impossible that these animals could drift on a log for the forty miles or more required for transportation to the outer islands. The Pleistocene elephants of Santa Cruz Island are undoubted evidence of a former land connection at this place.

The fringing islands of the West and East Indies and of such groups as Madeira and Cape Verde Islands of the Atlantic have separations of intermediate depths, and like the California islands have faunal evidence of former connections. Extensive bibliographies cover this subject but the evidence is not entirely conclusive.

SUMMARY AND CONCLUSIONS

An examination of available information from the sea floor around the continents and around oceanic islands provides abundant evidence of large scale submergence. The investigation of submarine canyons has now yielded so much indication of subaerial origin that the world-wide distribution of these features constitutes the most conclusive evidence that the continental margins and even some of the oceanic islands once stood thousands of feet higher in relation to sea level than they do at the present time. Widespread submergence of oceanic islands in the coral island areas, long ago advocated by Darwin, has now been well confirmed by recent borings in widely separated places. This evidence indicates that extensive tracts of the ocean floor have been considerably depressed. The recent discoveries of flat platforms on the tops of deep seamounts suggest a former elevated condition during which waves bevelled the platforms to sea level, or that the flat surface was developed at sea level by calcareous reef builders. Evidence from bottom samples is somewhat less conclusive than that of the canyons and coral islands, but the limited dredgings and corings available from deep water contain a large per cent which have evidence of former shallow-water conditions. Insular faunas in general are also of a character which favors considerable submergence, particularly along the continental margins. The deep fills in many of the mouths of river valleys, not discussed here, are still further evidence in the same direction.

---

[2] A term referring to broad submerged areas seaward of and deeper than the continental shelves but landward of the principal continental slope.

The significance of all of these converging lines of evidence in the unravelling of the late geological history should be profound. Unless further exploration provides very different evidence from that which is now available, we shall be forced to conclude that the continental margins and the islands have undergone great submergence alike in areas where stability was supposed to have long prevailed and in areas where other lines of evidence indicate either great crustal rising (as along much of the California Coast) or long-continued crustal submergence (as at the delta of the Mississippi). The world-wide nature of the evidence suggests a change of sea level, and the indication of Pleistocene origin of at least some of the submarine canyons could be used as an indication that the changing sea level resulted from the extraction of water during glaciation. However, the extent to which the sea could be lowered by this means is most difficult to determine. The common estimates of 40 fathoms may be far too small (Shepard and Emery, 1941: 146–155), since the ice could have been much thicker on the continents than generally assumed and may have covered much more territory than is considered in the usual estimates, particularly in the Arctic and along the continental margins. On the other hand, some of the recent evidence indicates that the submergence, at least of the coral reef islands, was going on all through the Tertiary. Possibly the same may be true of the submarine canyons, and the evidence of Pleistocene origin now available may apply only to the upper nearshore portions where glacially-controlled sea levels may have been important factors. However, the failure to find breaks in the canyons between any apparent Pleistocene excavation and canyons of earlier age is disconcerting. None of the well-explored canyons which can be traced into shallow water show a change in character at a depth of the order of 40 fathoms (the supposed glacial low sea level), nor has any other depth down to about 1,000 fathoms shown any indication of such a change.

In conclusion we are confronted with abundant evidence of world-wide submergence, some of it Pleistocene and some Tertiary or even earlier. It is time that these facts be faced by those geologists who have been trying to hold back the tide. More exploration is badly needed, but new hypotheses should utilize the wealth of oceanographic data now available.

REFERENCES

BEEBE, C, WILLIAM

1932. The world of Bermuda fish. Bull. N. Y. zool. Soc., *35* (2): 36–42.

BUCHER, W. H.

1940. Submarine valleys and related geologic problems of the North Atlantic. Bull. geol. Soc. Amer., *51:* 489–512.

DALY, R. A.

1910. Pleistocene glaciation and the coral reef problem. Amer. J. Sci., (4) *30:* 297–308.

1936. Origin of submarine canyons. Amer. J. Sci., (5) *31:* 401–420.

DANGEARD, LOUIS

1928. Observations de géologie sous-marine et d'océanographie relatives à la Manche. Ann. Inst. océanogr. Monaco, *6* (1): 1–295.

DARWIN, CHARLES

1842. Selected works of Charles Darwin. The structure and distribution of coral reefs. D. Appleton and Co., New York and London. Vol. *13*. 366 pp.

DAVIS, W. M.

1928. The coral reef problem. Spec. Publ. Amer. geogr. Soc., 9: 1–596.

DOBRIN, M. B., B. L. SNAVELY, G. WHITE, R. BERESFORD AND B. PERKINS, JR.

1946. Seismic-refraction survey of Bikini Atoll. Abstract in: Bull. geol. Soc. Amer., *57* (2): 1189.

FISK, H. N.

1944. Geological investigation of the Alluvial Valley of the lower Mississippi River. War Dept., Corps of Engineers, U. S. Army. 78 pp.

HANZAWA, SHOSHIRO

1940. Micropaleontological studies of drill cores from a deep well in Kita-Daito-Zima (North Borodino Island). Jubilee Publ., Prof. H. Yabe, *2:* 755–802.

HESS, H. H.

1946. Drowned ancient islands of the Pacific basin. Amer. J. Sci., *244:* 772–791.

HULL, EDWARD

1912. Monograph on the sub-oceanic physiography of North Atlantic Ocean. Edward Stanford, London. 41 pp.

JOHNSON, D. W.

1939. The origin of submarine canyons. Columbia Univ. Press, New York. 126 pp.

KUENEN, PH. H.

1933. Geology of coral reefs. Geol. Res. Snellius Exped., 1929–1930, *5* (2): 1–125.

1947. Two problems of marine geology: atolls and canyons. Verh. Akad. Wet., Amst., Natuurk., (Tweede Sec.) *43* (3): 1–68.

LADD, H. S., J. I. TRACEY, AND G. G. LILL

1948. Drilling on Bikini Atoll, Marshall Islands. Science, *107* (2768): 51–55.

NEEB, I. G. A.

1943. Bottom samples. The composition and distribution of the samples. Geol. Res. Snellius Exped., 1929–1930, *5* (3) sec. 2: 55–268.

Pettersson, Hans

1945. Iron and manganese on the ocean floor. Goteborg. VetenskSamh. Handl., Följ. 6, Ser. B, Bd 3, no. 8: 1–37.

1946. Oceanographic work in the Mediterranean. Geogr. J., *107:* 163–166.

Richards, H. C. and Dorothy Hill

1942. Great Barrier Reef bores, 1926 and 1937: Descriptions, analyses, and interpretations. Rep. Great Barrier Reef Exped., *5:* 1–122.

Shepard, F. P.

1941. Nondepositional physiographic environments off the California Coast. Bull. geol Soc. Amer., *52* (2): 1869–1886.

1948. Submarine geology. Harper and Bros., New York. 338 pp.

Shepard, F. P. and K. O. Emery

1941. Submarine topography off the California Coast. Spec. Pap. geol. Soc. Amer., 31: 1–171.

Trask, P. D., F. B. Phleger, Jr. and H. C. Stetson

1947. Recent changes in sedimentation in the Gulf of Mexico. Science, *106* (2759): 460–461.

Wegener, A. L.

1922. The origin of continents and oceans (Trans. by J. G. A. Skerl). E. P. Dutton & Co., New York. 205 pp.

OCEANOGRAPHY AND GEOSYNCLINES[1]

By

PARKER D. TRASK

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts

Geosynclines are one of the fundamental problems of geology. Since they are intimately related to the ocean or some inland body of marine water, knowledge of modern concepts of oceanography helps the geologist in understanding better their mode of origin. Geosynclines are masses of sediment, largely clastic in nature, 30,000 to 50,000 feet thick.[2] The deposits consist of alternating sandstone, siltstone, and shale.

The presence of sandstone, together with such features as ripple marks and crossbedding, has suggested varying conditions of deposition in shallow water. Thus the rate of settling of the sea bottom has been considered as being approximately equal to the rate of accumulation of sediment. The alternation between coarse and fine sediments has been taken as evidence of rhythmic changes in (1) sea level, (2) rate of sinking, or (3) rate of deposition, or a combination of these three factors. In the orthodox concept, sedimentation continues until the whole mass of sediment is upturned into a mountain range.

No uniformity of opinion has been reached as to how thick an accumulation of sediment is required before it is called a geosyncline, but the common figure is 30,000 feet. Although many people regard the load of sediment as instrumental in the settling of the basin, and though all are agreed that independent diastrophic forces aid the process, there is no unanimity of opinion as to the relative effects of load and diastrophism. Nevertheless many people believe that the sediments are deposited in essentially shallow water. The presence of sandstone, crossbedding, ripple marks, foreset beds, rain prints, peat, and shallow-water forms of life are evidence for this inference. However, recent work in submarine geology indicates that some if not a significant proportion of geosynclinal sediments were deposited in deep water.

Deposition of sediments in the sea is influenced materially by the physical conditions, that is, the dynamics of the water. Sverdrup, in recent years, has led the way in our understanding of the physical processes of the sea, and any geologist who is interested in learning some of these principles is referred to his works, particularly "The Oceans" (Sverdrup, *et al.*, 1942).

---

[1] Contribution No. 454 from the Woods Hole Oceanographic Institution.

[2] For a general study of geosynclines, see Glaessner and Teichert, 1947.

The principal oceanographic factors that govern deposition in geosynclines are currents, waves, wind, and configuration of the sea bottom, and to a minor extent, ecological considerations affecting the growth of organisms, including temperature and salinity of the water. To these should be added the typically geologic factors—supply of detritus and diastrophism.

Currents are of four principal types: oceanic, wind, longshore, and tidal. The oceanic currents, such as the Gulf Stream and the Japanese Current, are largely products of the open sea and are not likely to have great influence upon geosynclinal deposition, though offshoots may enter inland basins or embayments of deposition. Due to the Coriolis force, these currents in the northern hemisphere flow northward along the west side of oceans and southward along the east side. Similarly, wind-driven currents in the northern hemisphere trend in a direction to the right of the wind, and thus on the average are likely to flow northward on the west side of basins and southward on the east side, but of course such directions depend upon the direction of the prevailing wind. As both wind and oceanic currents tend to have their maximum velocities offshore, they do not materially affect deposition of sediments until the detritus has been transported some distance from the coast. Tidal currents have their greatest influence in lagoons and estuaries. In some bodies such as the Gulf of Mexico they have little effect.

Longshore currents on the other hand have a pronounced effect upon geosynclinal deposition. As Sverdrup and his co-workers (Putnam, *et al.*, in press) have shown, longshore currents commonly are formed by waves that approach the shore at an angle, with their course along the beach in the general direction with which the waves approach the beach. Since the waves are formed by the wind, the prevailing wind has a decided effect upon the transport of detritus along the shore. The proportion of fine particles in the detritus influences materially the type of deposition. If the proportion of fine material is large, currents may transport the bulk of the sediment a long distance from the mouth of the streams before it is deposited. If it is small, most of the sediment is likely to be laid down near the mouths of the streams, unless the longshore currents are strong.

Currents may extend with varying velocity to the bottom of the sea. Irregularities on the sea floor interfere with the free movement of the currents and may set up eddies which interfere with deposition.

Consequently, sediments over humps or ridges on the sea floor, or even along slopes that are relatively steeper than adjacent parts of the sea bottom, are likely to be coarser than deposits in adjoining areas. In as much as the effect of shape of the sea bottom is independent of depth of water, coarse sediments may be found in water of any depth.

Waves influence deposition by stirring up the bottom in shallow water, thus effecting a winnowing action, which tends to cause sediments on flat continental shelves to be relatively coarse, even at considerable distance from shore and in water up to 100 fathoms in depth. Waves also wear back the coast lines and thus supply detritus. Their effect upon the profile of equilibrium of the sea floor is not yet well understood, but it may take a very long time for the so-called profile of equilibrium to be reached, during which interval of time sea level may change, thus causing the ocean to seek a new level of equilibrium.

In recent years oceanographers have found that internal waves, which never reach the surface, are more common than was formerly thought. Active research is in progress, which may indicate that internal waves are capable of generating currents that affect the transport and winnowing of sediments at considerable distances from shore and in great depths of water.

The wind, as mentioned above, profoundly affects the waves, and in turn the currents. It also influences the climate and ecologic conditions of sedimentation.

The configuration of the basin of deposition probably has a greater effect on geosynclinal deposition than has been generally recognized in the past, because of its pronounced influence on currents, both along shore and in deep water.

The relationship of ecology to geosynclinal deposition is too complicated to discuss here, but the temperature of the water may have a profound effect. For example, if the water is very cold, settling velocities of detritus are lessened, with the result that sediment will be transported greater distances before it is deposited. Similarly, the mixing of fresh and salt water at river mouths has decided effects upon coagulation and settling of clay particles.

With respect to supply of detritus, pertinent factors are: rate of supply, grain size, strength and constancy of the currents, and configuration of the basin adjacent to the rivers that supply the debris. Even though a river carries a heavy load, if all the particles are fine-grained, currents may remove the detritus almost as fast as it arrives. The mode of origin of bars and beaches is understood by geologists, but the formation of submerged bars or underwater dunes is less well understood. The latter have been observed in several places in recent years, even at some distance from shore, and they may be more common than has been recognized.

Much is spoken about foreset beds upon deltas. The formation of foreset beds depends upon the development of an essentially flat but gently sloping surface of deposition terminated by a more abrupt slope, down which detrital particles can roll after they have been pushed off the topset part of the delta. The angle of inclination of the foreset beds depends to some extent upon the depth of water at the forward edge of the topset beds. In glacial lakes or fjords, where the water level may have risen very rapidly with respect to rate of delta formation, the detritus may be dumped into relatively deep water, with the result that the classic type of steep foreset beds may form. But in the sea where the rate of change in level of the water in general is much slower than in glacial lakes, the depth of water at the front of the topset beds may be relatively shallow, with the result that the angle of inclination of the foreset beds may be comparatively small, perhaps only one or two degrees. Consequently steep foreset beds are not likely to be characteristic of marine deltas, except of course in lagoons and sloughs along delta distributaries.

Diastrophism, as has been pointed out by many geologists, has a dominant influence on geosynclinal deposition. It affects the supply of detritus by governing the gradient of streams and by influencing wind, temperature and rainfall. It affects the deposition of debris by changing the configuration of the sea floor. Fluctuations in rate of diastrophism may lead to rhythmic alternations of sediments.

Much of what has been said above is familiar to most geologists, but perhaps certain aspects have not been emphasized in the past as much as they deserve. Among such aspects are: (1) The relationship between supply of detritus and longshore currents, (2) the effect of wind upon sedimentation, (3) the relationship of configuration of the basin to deposition, (4) the existence of relatively strong currents in deep water, which lead to deposition of coarse sediments far from shore, (5) the lack of steep foreset beds on large deltas and, (6) material differences in rate of settling with respect to rate of deposition.

In order for sediments to accumulate in shallow water, these two rates must be essentially the same. If they are the same, either the influx of debris fortuitously balances the rate of sinking or the influx of debris controls the rate of settling. Natland (1933) has shown that the succession of foraminifera in the present Channel Islands region of California from deep water to shallow water is essentially the same as in the Lower Pliocene formations in the near-by Los Angeles basin. This relationship suggests that the lower beds in the Lower Pliocene were deposited in deep water and that the basin filled up with relatively little additional sinking. A similar sequence of foraminifera has been reported from the Lower Pliocene in the Eel River basin in California. Thus it would seem that, at first, downwarping was greater than rate of sedimentation, following which sedimentation was greater than settling. As Barrell (1913, 1914) has pointed out for the Devonian geosyncline in Pennsylvania, the basin eventually filled up and continental beds were deposited, thus pushing the shoreline seaward. Hence deposition was more rapid than settlement.

Recent work by Trask, Phleger and Stetson (1947) in the Gulf of Mexico has confirmed this concept. In the central part of the Gulf, in water more than 11,000 feet deep, foraminiferal ooze is underlain at depths of 2 to 10 feet by alternating clay, silt, and sand containing a distinctly cold water fauna, characteristic of the waning stages of the Ice Age. The sand and silt is extremely well-sorted. In two places crossbedding or ripple marks were found in coarse clastics of the Ice Age deposits; in two other places, well-sorted sand zones, one and three feet thick respectively, were encountered at depths of more than three feet beneath the surface of the sediments. Such deposits, if hardened into rock and found in a geosyncline, would be taken as compatible with the idea of shallow-water deposition. Yet they were encountered in 11,000 feet of water.

If these sediments are of shallow-water origin, then within the last 20,000 to 100,000 years either the water of the Gulf of Mexico has been lowered more than 10,000 feet, or the sea bottom has been depressed that amount. Either idea is contrary to prevalent concepts. Of course, the continental glaciers might have been much thicker than they have previously been considered, with the result that 10,000 feet of water might have been removed from the sea. Such an idea is in accord with the subaerial origin of the many submarine canyons that have now been found. The low level might have prevailed for only a few thousand years, since most submarine canyons do not have deep or wide gorges. Such a concept, however, is contrary to prevailing opinion.

Similarly, if the sea floor has been lowered, the rate of settling would be one foot in 2 to 10 years—truly a rapid rate. Also, if the sea floor settled 10,000 feet, where did the mass of rock beneath the floor of the Gulf go? No corresponding recent rise of sea level in adjacent lands has been observed, unless one wishes to ascribe some of the uplift of the Mexican Plateau to recent downwarping of the Gulf. Such an idea of course is most unorthodox.

The concept of deposition in shallow water thus seeming to be anomalous, the only alternative is deposition of the sediments in deep water. If the well-sorted sand, three feet thick, was deposited in

11,000 feet of water, as were also the crossbedded sands and the numerous well-sorted silt layers, then the water must have been in motion to account for the high degree of winnowing exhibited by the sediments. Perhaps internal waves caused such motion. At any rate, the presence of such phenomena, ordinarily ascribed to shallow water, is most peculiar. Geologists might do well to exercise caution when ascribing a shallow-water origin to sediments upon such evidence as well-sorted sandstone, crossbedding, or ripple marks.

SUMMARY

In conclusion, the data presented in this paper support the inference that deposition in a geosyncline is a function of the ratio of rate of deposition to the rate of settling. If the rate of settling is greater, the sediments may accumulate in deep water; if the rate of deposition is greater, particularly in the later stages of the formation of the geosyncline, the sediments may accumulate in shallow water or upon land. Mathematically, the sequence of events may be expressed by the following equation:

$$ds/dt - dd/dt = f\,(D)\,,$$

where ds/dt is the rate of settling, dd/dt is the rate of deposition, and $f\,(D)$ is a composite function representing the combined effects of diastrophism throughout the areas of loading and unloading of the sediments, together with the effect of diastrophism upon climate.

Perhaps such a mathematical formula is of little meaning at present, but ultimately, as geology, like every other science, becomes more quantitative, tangible values can be substituted in the equation. When that day comes geologists will be able to interpret with satisfaction the mode of origin of the geosynclines they study.

REFERENCES

BARRELL, JOSEPH
1913. The Upper Devonian delta of the Appalachian geosyncline. Amer. J. Sci., (4) *36:* 429–472; 1914, (4) *37:* 87–109 and 225–253.

GLAESSNER, M. F. AND C. TEICHERT
1947. Geosynclines: a fundamental concept in geology. Amer. J. Sci., *245:* 465–482 and 571–591.

NATLAND, M. L.
1933. The temperature- and depth-distribution of some recent and fossil foraminifera in the southern California region. Bull. Scripps Inst. Oceanogr., Tech. Ser. *3:* 225–230.

PUTNAM, J. A., W. H. MUNK AND M. A. TRAYLOR
Longshore currents. Trans. Amer. geophys. Un.

SVERDRUP, H. U., M. W. JOHNSON AND R. H. FLEMING
1942. The oceans. Prentice-Hall, Inc., New York. 1087 pp.

TRASK, P. D., F. B. PHLEGER, JR. AND H. C. STETSON
1947. Recent changes in sedimentation in the Gulf of Mexico. Science, *106* (2759): 460–461.

ERRATA

JOURNAL OF MARINE RESEARCH

(Volume VII, Number 2)

Page 77, footnote 2 should read:

[2] Terms of the form $\frac{\bar{\tau}_x(y) - \tau_x(x_1y)}{\Delta x} \cdot \frac{dx_1}{dy}$, thereby neglected, are of small conse-

Page 78, line 1 should read:

by x_2; $\Delta P = P(x_1y) - P(x_2y)$ is the difference of P for two stations at

Page 82, Table I, last line should read:

8..........17.2-20.7..........1895.........10.80.........10.85

Page 92, Equation (22) should read:

$$\frac{\partial P}{\partial y} = \frac{\Delta P}{\Delta y} - \frac{\partial P}{\partial x}\frac{1}{\tan\beta}, \tag{22}$$

Page 93, Equation (25) should read:

$$\begin{aligned}&+\frac{bnK}{(k^2-K^2)}\Big[k\cos 2k(\varphi-\varphi_0)\cos 2K(\varphi-\varphi_0'')\\ &+K\sin 2k(\varphi-\varphi_0)\sin 2K(\varphi-\varphi_0'')\Big],\end{aligned} \tag{25}$$

Page 97, line preceding Equation (29) should read:

(1a) and (6):

NOTE TO AUTHORS

The pages of this Journal are open to all contributors, and manuscripts may be submitted to any member of the Board of Editors or to the Managing Editor. Publication will be made only in the English language, but copy may be submitted for consideration in other languages. Aid in translation of accepted manuscripts will be given when possible.

All manuscripts should be typewritten with double space, carefully corrected, with footnotes continuously numbered and set off in the body of the text. If an article is to appear as a contribution from another institution, it should be approved by that institution before submission; contribution numbers of such institutions are submitted with the return of galley proofs.

Each author receives fifty reprints (with covers) free of charge up to a maximum of 150 free reprints of any single article in the case of joint authorship by three or more authors. Additional reprints with or without covers may be ordered by authors or institutions at the cost of printing. Orders for reprints are given when galley proofs are returned.

Both galley and page proofs are submitted to the senior author of each article.

Tables should be prepared for clarity and as near as practicable to Journal pattern, without ruled lines if possible. Table footnotes should appear at the base of each Table.

Figures are required for consideration simultaneously with the manuscript. Original figures of inconvenient size for forwarding may be represented by photographs on first submission of a manuscript; originals are required for line cuts. Illustrations requiring half-tone reproduction should be avoided as far as possible, but they will be accepted when necessary. *It is very important that the proper size of characters and symbols should be used for the lettering of illustrations.* A diagram for determining suitable heights of letters for original drawings that will be reduced, together with an explanatory note, has been published in this Journal, Vol. III, No. 1: 89–91; reprints will be submitted to authors on request. Mechanical lettering is recommended. If facilities for lettering are not available to the author, it is preferable to receive illustrations with lettering indicated in pencil only, to be finished in ink in the editorial office. All illustrations will be subject to at least a minimum reduction at the rate of 1 : 1.2. Drawings which have been heavily patched and pasted must *not be rolled* for mailing. The legends for the figures, identified by corresponding figure numbers on the illustrations also, should be typed in sequence at the end of the manuscript.

A complete and *accurate list of references* in alphabetical order at the end of each article is essential. References may be indicated in the text by author's name (or names) and *year of publication*, or by italicized numbers which correspond to the numbers given in sequence in the list of references. Each reference should include the last name followed by at least two initials or first name; co-authors should be listed with the initials or first name preceding the last name. The date should be indented on the line below, followed by a complete title *verbatim* in lower case except for the first word. Publisher, city and pagination should be indicated for books; for periodicals (abbreviations from "World List of Scientific Periodicals," Oxford, as far as possible), the volume (italicized), number and pagination should be given.

Additional information regarding the preparation of manuscripts will be furnished by the Managing Editor.